Алгебра. Конспект лекций

Иваньшин П.Н.

2 мая 2012 г.

Оглавление

Глава 1

Множества

Определение 1. *Разбиение множества S — такое множество π его подмножеств, что*

a) Если $A \in \pi$, то $A \neq \emptyset$.

b) Если $A \in \pi$ и $B \in \pi$, то либо $A = B$, либо $A \bigcap B = \emptyset$

c) каждый элемент множества S принадлежитнекоторому элементу множества π.

То есть, разбиение множества S — семейство его непустых подмножеств, таких, что каждый элемент из S принадлежит в точности одному подмножеству из этого семейства.

Пример 1. *Пусть $S = \{1, 2, 3, 4, 5\}$. Тогда $\pi = \{\{1, 2, 3\}, \{4, 5\}\}$ — разбиение S на два подмножества.*

1.1 Отношения эквивалентности

Определение 2. *Бинарное отношение R на непустом множестве A — подмножество декарова произведения $A \times A$, то есть $R \subset A \times A$.*

Пример 2. *Отношение "меньше или равно" на множестве $\mathbb{R}$ можно задать подмножеством $\mathbb{R} \times \mathbb{R} \supset R = \{(x, y)|y - x \geq 0\}$.*

Можно определить и отношение, обратное к данному, как $R^{-1} = \{(y, x)|(x, y) \in R\}$.

Определение 3. *Отношение эквивалентности E на множестве A — бинарное отношение $E \subset A \times A$, удовлетворяющее условиям*

a) Рефлексивность: $\forall x \in A \ (x, x) \in E$.

b) Симметричность: $\forall (x, y) \in E \ (y, x) \in E$.

c) Транзитивность: $(x, y), (y, z) \in E$ влечет $(x, z) \in E$.

Будем обозначать через $x \equiv_E y$ или $x \equiv y$ (x эквивалентно y), если $(x, y) \in E$.

Класс эквивалентности элемента $x \in A$ — множество $[x] = \{y \in A | (x, y) \in E\}$. Любой элемент $y \in [x]$ называется представителем класса эквивалентности $[x]$.

Задача 1. *Доказать, что* $y \in [x] \Leftrightarrow [y] = [x]$

Определим, наконец, множество всех классов эквивалентности, или так называемое фактормножество $A/\equiv_E = \{[x] | x \in A\}$.

Пример 3. *Пусть* $A = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$. *Определим отношение эквивалентности на* A *следующим образом;* $(x, y) \equiv (x', y') \Leftrightarrow xy' = x'y$. *Проверить, что* $\equiv$ *— отношение эквивалентности.*

Тогда $A/\equiv \cong \mathbb{Q}$.

Предложение 1. *Если* E *— отношение эквивалентности на множестве* A, *то фактор-множество* $A/E = \{[a] | a \in A\}$ *— разбиение множества* A. *И наоборот, если* π *— разбиение* A, *то существует такое отошение эквивалентности* E *на* A, *что* $\pi \cong A/E$.

Доказательство — упражнение.

1.2 Теорема о факторизации

Пусть $f : A \to B$. Определим отношение эквивалентности на множестве A по правилу: $x \equiv x'$, если $f(x) = f(x')$.

Задача 2. *Доказать, что* $\equiv$ *— отношение эквивалентности.*

Определим $i : A/\equiv \to B$ равенством $i([x]) := f(x)$. Пусть $s : A \to A/\equiv$ — каноническое отображение на фактормножество.

Теорема 1. *Пусть* $f : A \to B$ *— произвольная функция. Тогда диаграмма коммутативна, то есть* $f = i \circ s$.

Задача 3. *Доказать теорему.*

Глава 2

Группы

Определение 4. *Группа — множество G с бинарной операцией $*$: $G \times G \to G$, удовлетворяющее следующим свойствам:*

a) $\forall a, b \in G$, $a * b \in G$.

b) Ассоциативность: $\forall a, b, c \in G$, $a * (b * c) = (a * b) * c$.

c) Существование нейтрального элемента: $\exists e \in G$, $\forall a \in G$, $a * e = e * a = a$.

d) Существование обратного элемента: $\forall a \in G\ \exists a^{-1} \in G$, $aa^{-1} = a^{-1} * a = e$.

Если опустить последнее требование, то получим определение полугруппы.

Определение 5. *Пусть* $(G_i, *_i)$, $i = 1.2$ *— группы. Отображение* $f : (G_1, *_1) \to (G_2, *_2)$ *— гомоморфизм групп, если* $\forall a, b \in G_1$ $f(a *_1 b) = f(a) *_2 f(b)$.

Если f — биекция G_1 на G_2, то f — изоморфизм.

Задачи

1. Ассоциативна ли операция $*$ на множестве M, если

a) $M = \mathbb{N}$, $x * y = x^y$;

b) $M = \mathbb{N}$, $x * y = \mathrm{GCD}\{x, y\}$;

c) $M = \mathbb{N}$, $x * y = 2xy$;

d) $M = \mathbb{Z}$, $x * y = x - y$;

e) $M = \mathbb{Z}$, $x * y = x^2 + y^2$;

f) $M = \mathbb{R}$, $x * y = x/y$;

2. Доказать, что во всякой конечной полугруппе найдется идемпотент (то есть $\exists x\ inG$, $\exists n \in \mathbb{N}$, $x^n = x$)

3. Полугруппа моногенна, если состоит из положительных степеней одного из своих элементов, который называется порождающим. Доказать, что

a) моногенная полугруппа конечна $\Leftrightarrow$ содержит идемпотент;

b) конечная моногенная полугруппа либо группа, либо имеет только один порождающий элемент;

c) любые две бесконечные полугруппы изоморфны.

4. Какие из указанных множеств с операциями есть группы:

a) $(A, +)$, где A одно из множеств $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{Q}$;

b) $(A, \bullet)$, где A одно из множеств $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{Q}$;

c) $(A^*, \bullet)$, где A одно из множеств $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{Q}$, $A^* = A \setminus \{0\}$;

d) $(n\mathbb{Z}, +)$, $n \in \mathbb{N}$.

e) $(\{-1, 1\}, \bullet)$.

5. Доказать, что коммутатор $[x, y] = xyx^{-1}y^{-1}$ элементов x, y группы G обладает свойствами

a) $[x, y]^{-1} = [y, x]$;

b) $[xy, z] = x[y, z]x^{-1}[x, z]$;

c) $[z, xy] = [z, x]x[z, y]x^{-1}$.

2.1 Кольца и поля

Определение 6. *Кольцо $(R, +, \cdot)$ — алгебраическая система, удовлетворяющая условиям:*

a) R — коммутативная группа относительно операции $+$, (то есть $\forall x, y \in R$, $x + y = y + x$). Единичный элемент 0_R называется нулевым.

b) R — полугруппа с единицей 1_R по умножению.

c) Дистрибутивность: $\forall x, y, z \in R$ $x(y + z) = xy + xz$ и $(y + z)x = yx + zx$.

Пример 4. *Множества $\mathbb{Z}, \mathbb{R}, \mathbb{Q}$ — кольца относительно обычных операций.*

Кольцо R называется коммутативным, если $\forall a, b \in R$, $ab = ba$. Если $\forall n \in \mathbb{N}$, $\sum\limits_{i=1}^{n} 1_R \neq 0$, то говорят, что R имеет характеристику 0.

Определение 7. *Элемент $a \in R$ называется делителем нуля, если $\exists b \in R$, $b \neq 0$, $ab = 0$. Элемент $u \in R$ обратим, если $\exists u^{-1} \in R$, $uu^{-1} = 1$.*

Область целостности — нетривиальное коммутативное кольцо без делителей нуля.

Пример 5. *Кольцо $\mathbb{Z}$ — область целостности.*

Определение 8. *Поле — нетривиальное коммутативное кольцо, в котором каждый ненулевой элемент обратим.*

Пример 6. *Множества* $\mathbb{R}$ *и* $\mathbb{Q}$ *со стандартными арифметическими операциями — поля.*

Задачи

1. Какие из следующих множеств образуют кольцо относительно операций сложения и умножения:

a) $\mathbb{Z}$;

b) $n\mathbb{Z}$;

c) $\mathbb{Z}^+ \bigcup \{0\}$;

d) $\mathbb{Q}$;

e) $\{x + \sqrt{2}y | x, y \in \mathbb{Q}\}$;

f) $\{x + \sqrt[3]{2}y | x, y \in \mathbb{Q}\}$;

2. Какие из приведенных выше колец содержат делители нуля?

3. Пусть R — кольцо с единицей, $x, y \in R$. Доказать, что

a) Обратимость xy и yx влечет обратимость x и y.

b) Если R без делителей нуля, то обратимость xy влечет обратимость x и y.

2.2 Подгруппы

Пусть G — группа.

Определение 9. *Непустое подмножество* H *в группе* G *называется подгруппой, если вместе с любыми двумя его элементами оно содержит их произведение, и с каждым своим элементом* H *содержит его обратный.*

Предложение 2. *Если* H *— подгруппа в группе* G *и* e *— единичный элемент* G*, то* $e \in H$*.*

Задача 4. *В произвольной группе произведение любого числа элементов не зависит от расстановки скобок.*

Предложение 3. *Для непустого подмножества* H *в группе* G *следующие условия эквивалентны:*

A) H *является подгруппой в* G*;*

B) если $x, y \in H$*, то* $xy^{-1} \in H$*.*

Доказательство. Пусть выполнено условие A), и $x, y \in H$. В силу определения 9 получаем $x, y^{-1} \in H$, откуда $xy^{-1} \in H$, т. е. выполнено условие B). Обратно, пусть выполнено условие B), и $y \in H$. Тогда $y, y^{-1} \in H$, откуда $e = yy^{-1} \in H$ по B). Далее $e, y \in H$, откуда $y^{-1} = ey^{-1} \in H$ по B). Наконец, если $x, y \in H$, то $x, y^{-1} \in H$ по доказанному выше. Отсюда $x(y^{-1})^{-1} = xy \in H$. □

Задача 5. *Если H_i,$i \in I$ — подгруппы группы G, то $\bigcap_{i \in I} H_i$ — подгруппа группы G.*

Определение 10. *Пусть $a \in G$. Для произвольного целого числа n положим*

$$a^n = \begin{cases} e & n = 0 \\ a \cdots a & n > 0 \\ a^{-1} \cdots a^{-1} & n < 0 \end{cases}$$

Предложение 4. *Пусть a — элемент некоторой группы и $n, m \in \mathbb{Z}$. Тогда $a^{n+m} = a^n a^m$, $(a^n)^m = a^m n$.*

Определение 11. *Пусть a — элемент некоторой группы. Порядком $|a|$ элемента a называется такое наименьшее натуральное число n, что $a^n = e$. Если такого числа n нет, то говорят, что порядок a равен бесконечности.*

Предложение 5. *Пусть $|a| = n \leq \infty$, и $m \in \mathbb{Z}$. Следующие условия эквивалентны:*

A) $n|m$ (n делит m);

B) $a^m = e$.

Определение 12. *Пусть $a \in G$. Через (a) обозначим множество $\{a^n | n \in \mathbb{Z}\}$ всех степеней элемента a.*

Задача 6. *(a) является подгруппой в G.*

Определение 13. *Пусть G — группа, а S — подмножество G. Говорят, что S порождает G, или, что S — семейство генераторов G, если $\forall g \in G$, $\exists x_1, \ldots, x_n \in S$, $(g = x_1^{e_1} \cdots x_n^{e_n}$, $e_i = \pm 1)$.*

Задача 7. *Доказать, что семейство всех таких произведений — 1) подгруппа G, 2) наименьшая подгруппа G, содержащая S.*

Пример 7. *Существует две неабелевы группы порядка 8.*

Одна — группа симметрий квадрата, порожденная такими двумя элементами σ и τ, что $\sigma^4 = \tau^2 = e$ и $\tau\sigma\tau^{-1} = \sigma^3$.

Вторая — группа кватернионов, порожденная двумя элементами i и j, такими, что если ввести в рассмотрение еще два элемента $k = ij$ и $m = i^2$, получим соотношения $i^4 = j^4 = k^4 = e$, $i^2 = j^2 = k^2 = m$, $ij = mji$.

2.3 Классы смежности

Пусть G — группа, а H — подгруппа G.

Определение 14. *Правый класс смежности по H в G — подмножество G вида aH для некоторого элемента $a \in G$. Любой элемент aH называется представителем класса смежности aH.*

Задача 8. *Отображение $h \mapsto ah$ — биекция H на aH.*

Следовательно, любые два класса смежности состоят из одого и того же количества элементов.

Предложение 6. *Пусть $a, b \in G$ и $aH \bigcap bH \neq \emptyset$. Тогда $aH = bH$.*

Доказательство. Пусть $ax = by$ $x, y \in H$. Тогда $a = byx^{-1}$. Но $yx^{-l} \in H$. ледовательно, $aH = b(yx^{-1})H = bH$, так как $\forall z \in H$ $zH = H$. □

Таким образом, G — дизъюнктное объединение левых классов смежности по H. То же утверждение верно и для правых классов смежности (то есть подможеств G вида Ha). Обозначим число левых классов смежности в группе G по подгруппе H через $(G : H)$, и назовем (левым) индексом H в G. Таким образом, получаем

Предложение 7. *Пусть G — группа, а H — подгруппа G. Тогда $(G : H)(H : 1) = (G : 1)$, то есть, если два из рассматриваемых индеков конечны, конечен и третий, и тождество верно. Если $(G : 1)$ конечен, порядок H делит порядок G.*

Более общо, пусть H, K — подгруппы G и $H \supset K$. Пусть $\{x_i\}$ — семейство (левых) представителей K в H, а $\{y_j\}$ — семейство представителей H в G. Тогда $\{y_j x_i\}$ — семейство представителей K в G.

Доказательство. Заметим, что

$$H = \bigcup_i x_i K \text{ (дизъюнктное)}$$

$$G = \bigcup_j y_j H \text{ (дизъюнктное)}$$

Следовательно,

$$G = \bigcup_{j,i} y_j x_i K.$$

Покажем, что последнее объединение также дизъюнктно. Пусть $\exists i, j, i', j'$, $y_j x_i K = y_{j'} x_{i'} K$. Умножим слева на H и заметим, что $x_i, x_{i'} \in H$. Тогда $y_j H = y_{j'} H$, следовательно, $y_j = y_{j'}$. Тогда $x_i K = x_{i'} K$, то есть $x_i = x_{i'}$. □

Пример 8. *Группа простого порядка — циклическая. Пусть G порядка p, $a \in G$ и H — подгруппа G, порожденная a. Тогда порядок H делит порядок G, то есть, так как $a \neq 1$, $H = G$. Следовательно, группа G циклическая.*

2.4 Нормальные подгруппы

Пусть $f : G \to G'$ — гомоморфизм групп, и пусть H — его ядро. Если $x \in G$, то $xH = Hx$ так как оба множесьва есть $f^{-1}(f(x))$. Это же соотношение можно переписать как $xHx^{-1} = H$.

Пусть, напротив, G — группа, и H — ее подгруппа. Пусть $\forall x \in G$ $xH \subset Hx$ (или, что эквивалентно, $xHx^{-1} \subset H$). Если мы рассмтотрим x^{-1} вместо x, то получим $H \subset xHx^{-l}$, следовательно $xHx^{-1} = H$. Итак, наше условие эквивалентно условию $\forall x \in G$ $xHx^{-i} = H$.

Определение 15. *Подгруппа, удовлетворяющая этому условию, называется нормальной.*

Пусть G' — набор классов смежности H. (По предположению, правый смежный класс совпадает с левым, то есть можно опустить этот термин.) Если xH и yH — классы смежности, то их произведение $(xH)(yH)$ — также класс смежности, поскольку $xHyH = xyHH = xyH$. Таким образом, на G' определена бинарная операция, очевидно ассоциативная. Ясно, что единичным элементом для этой операции является класс смежности H, и что $x^{-l}H$ — обратный элемент для класса смежности xH. То есть, G' — группа.

Пусть $f : G \to G'$ — отображение, определенное по правилу $f(x)$ — класс смежности xH. Тогда f — гомоморфизм, и (подгруппа) H содержится в его ядре. Если $f(x) = H$, то $xH = H$. Так как H содержит единицу, $x \in H$. Следовательно, H совпадает с ядром гомоморфизма f, и верно утверждение, обратное приведенному в начале параграфа.

Группа классов смежности по нормальной подгруппе H обозначается через G/H. Отображение f G на G/H, построенное выше, называется каноническим отображением, и G/H называется факторгруппой G по H.

2.5 Идеалы колец

Определение 16. *Левым идеалом кольца A называется такое подмножество $\alpha \subset A$, что 1) α — подгруппа аддитивной группы A, 2) $A\alpha \subset A$*

(то есть, $A\alpha = \alpha$ поскольку $1 \in A$). Правый идеал определяется аналогично соотношением $\alpha A = \alpha$, а двусторонний идеал — множество. которое одновременно является и правым и левым идеалом. Двусторонний идеал часто называю просто идеалом.

Заметим, что 0 и A — сами идеалы A.

Если A — кольцо, и $a \in A$, то Aa — левый идеал, называемый главным. Говорят, что a — порождающий элемент (генератор) α (над A). Аналогично, AaA — главный двусторонний идеал, если мы определим AaA как набор всех сумм вида $\sum\limits_i x_i a y_i$, где $x_i, y_i \in A$. В более общем случае, пусть $a_l, \ldots, a_n$ — элементы A. Обозначим через $(a_l, \ldots, a_n)$ множество элементов A, которые можно записать в виде

$$x_1 a_1 + \ldots + x_n a_n, x_i \in A$$

Легко видеть, что это множество — левый идеал, а $a_l, \ldots, a_n$ называются генераторами левого идеала.

Если $\{\alpha_i\}_{i \in I}$ — семейство идеалов, то их пересечение $\bigcap\limits_{i \in I} \alpha_i$ — также идеал. Аналогичное утверждение верно для левых идеалов.

Задача 9. *Доказать, что если $\alpha = (a_l, \ldots, a_n)$, то alpha — пересечение всех левых идеалов, содержащих элементы $a_l, \ldots, a_n$.*

Определение 17. *Гомоморфизм колец — такое отображение $f : A \to B$, где A, B — кольца, что f — гомоморфизм, сохраняющий сложение и умножение на A и B. То есть, f удовлетворяет условиям: $f(a + a') = f(a) + f(a')$, $f(0_A) = 0_B$, $f(aa') = f(a)f(a')$, $f(1_A) = 1_B$.*

Ядром f называется его ядро как гомоморфизма групп по сложению.

Задача 10. *Ядро гомоморфизма колец $f : A \to B$ — идеал A.*

Пусть, наоборот, α — идеал кольца A. Построим факторкольцо A/α. Пусть A/α — факторгруппа группы A по сложению. Определим операцию умножения на A/α следующим образом: Пусть $x+\alpha$ и $y+\alpha$ — классы смежности по α, определим $(x+\alpha)(y+\alpha)$ как класс смежности $(xy+\alpha)$. Этот класс смежности корректно определен, так как если x_l, y_1 лежат в классах смежности x и y, соответственно, то и $x_1 y_1$ принадлежит классу смежности xy. Умножение тогда очевидно ассоциативно, относительно него определен единичный элемент, а именно, класс смежности $1+\alpha$, эта операция также дистрибутивна, поскольку дистрибутивность выполняется для представителей классов смежности. Таким образом, определена структура кольца на A/α, и каноническое отображение $f : A \to A/\alpha$ — гомоморфизм колец.

Задача 11. *1. Описать правые классы смежности при разложении группы G по подгруппе H.*

a) G — циклическая группа $\mathbb{Z}_8$ восьмого порядка, H — ее подгруппа четвертого порядка;

b) G — группа вращений куба, H — ее подгруппа, совмещающая с собой одну из граней куба;

c) G — группа всех невырожденных вещественных матриц, H — подгруппа матриц с определителем 1.

2. Доказать, что в конечной группе нечетного порядка любой элемент является квадратом другого, однозначно определенного, элемента.

3. Доказать, что если число правых классов смежности в разложении бесконечной группы G по подгруппе H конечно, то и число левых классов смежности конечно и равно числу правых классов.

4. Центром группы называется множество всех эле- элементов, коммутирующих со всеми элементами группы. Доказать, что центр есть нормальный делитель.

5. Пусть G — конечная группа, $H = \phi(G)$ — ее гомоморфный образ. Доказать, что порядок $x \in G$ делится на порядок $\phi(x) \in H$.

6. Коммутатором элементов a и b группы G называется $aba^{-1}b^{-1}$. Подгруппа, порожденная коммутаторами, называется коммутантом группы G. Доказать, что:

a) коммутатор a и b равен e тогда и только тогда, когда a и b коммутируют;

b) конечные произведения коммутаторов составляют коммутант;

c) коммутант является нормальным делителем;

d) факторгруппа по коммутанту абелева;

e) если G/H — абелева группа, то H содержит коммутант G.

7. Доказать, что если A и B — нормальные делители группы G и $a \in A$, $b \in B$, то $aba^{-1}b^{-1} \in A \bigcap B$.

Глава 3

Матрицы

Решение систем линейных уравнений.

$Ax = b$, где $A = (a_{ij})_{i,j=1}^{k,n}$, $x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_k \end{pmatrix}$

Определение 18. *Векторное (линейное) пространство или пространство строк длины n — множество $\mathbb{R}^n$ вместе с операциями сложения и умножения на скаляры (элементы $\mathbb{R}$), удовлетворяющее условиям:*

1) $X + Y = Y + X$ для любых векторов $X, Y \in \mathbb{R}^n$ (закон коммутативности);

2) $(X+Y)+Z = X+(Y+Z)$ для любых трех векторов $X, Y, Z \in \mathbb{R}^n$ (закон ассоциативности);

3) существует специальный (нулевой) вектор 0 такой, что $X+0 = X$ для всех $X \in \mathbb{R}^n$;

4) каждому $X \in \mathbb{R}^n$ отвечает противоположный (или обратный) вектор $-X$ такой, что $X + (-X) = 0$;

5) $1X = X$ для всех $X \in \mathbb{R}^n$;

6) $(\alpha\beta)X = \alpha(\beta X)$ для всех $\alpha, \beta \in mathbbR$ и $X \in \mathbb{R}^n$;

7) $(\alpha + \beta)X = \alpha X + \beta X$ для всех $\alpha, \beta \in mathbbR$ и $X \in \mathbb{R}^n$;

8) $\alpha(X + Y) = \alpha X + \alpha Y$ для всех $\alpha \in mathbbR$ и $X \in \mathbb{R}^n$.

Элементы векторного пространства называются векторами.

Система векторов $X_1, ..., X_k$ называется линейно зависимой, если найдутся k чисел $\alpha_1, \dots, \alpha_k$, одновременно не равных нулю и таких, что $\alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k = 0$.

Теорема 2. *Имеют место следующие утверждения:*

1) система векторов $\{X_1, \dots, X_k\}$ с линейно зависимой подсистемой сама линейно зависима;

2) любая часть линейно независимой системы векторов $\{X_1, \ldots, X_k\}$ *линейно независима;*

3) среди линейно зависимых векторов $X_1, \ldots, X_k$ *хотя бы один является линейной комбинацией остальных;*

4) если один из векторов $X_i, \ldots, X_k$ *выражается через остальные, то векторы* $X_1, \ldots, X_k$ *линейно зависимы;*

5) если векторы $X_1, \ldots, X_k$ *линейно независимы, а* $X_1, \ldots, X_k, X$ *линейно зависимы, то* X *— линейная комбинация векторов* $X_1, \ldots, X_k$*;*

6) если векторы $X_1, \ldots, X_k$ *линейно независимы и вектор* X_{k+1} *нельзя через них выразить, то система* $X_1, \ldots, X_k, X_{k+1}$ *линейно независима.*

Доказательство. 1) Пусть, например, первые s векторов $X_1, \ldots, X_s$, $s < k$, линейно зависимы, т.е. $\alpha_1 X_i + \ldots + \alpha_s X_s = 0$, где не все α_i равны нулю. Положив тогда $\alpha_{s+i} = \ldots = \alpha_k = 0$, получим нетривиальную линейную зависимость $\alpha_1 X_i + \ldots + \alpha_s X_s + \alpha_{s+1} X_i + \ldots + \alpha_k X_k = 0$.

Утверждение 2) непосредственно следует из i) (рассуждение от противного).

3) Пусть, например, $\alpha_k \neq 0$ в исследуемом соотношении. Тогда

$$X_k = -\frac{\alpha_1}{\alpha_k} X_1 - \ldots - \frac{\alpha_{k-1}}{\alpha_k} X_{k-1}.$$

Оставшиеся пункты — упражнение. □

Определение 19. *Пусть* V *— ненулевая линейная оболочка в* $\mathbb{R}^n$*, то есть линейное подпространство* $\mathbb{R}^n$*. Система векторов* $X_1, \ldots, X_r \in V$ *называется базисом для* V *(или в* V*), если она линейно независима и её линейная оболочка совпадает с* V*:*

$$\langle X_1, \ldots, X_r \rangle := \{\alpha_1 X_1 + \ldots + \alpha_r X_r | \alpha_i \in \mathbb{R}, i = 1, \ldots, r\} = V.$$

Лемма 1. *Пусть* V *— линейная оболочка в* $\mathbb{R}^n$ *с базисом* $X_1, \ldots, X_r$ *и* $Y_1, Y_2, \ldots, Y_s$ *— линейно независимая система векторов из* V*. Тогда* $s \leq r$*.*

Доказательство — упражнение. Указание — применить предыдущую теорему. В качестве простого следствия получаем

Теорема 3. *Каждая ненулевая линейная оболочка* $V \subset \mathbb{R}^n$ *обладает конечным базисом. Все базисы оболочки* V *состоят из одинакового числа* $r \leq n$ *векторов (это число называется размерностью оболочки* V *и обозначается* $\dim_{\mathbb{R}} V$ *или просто* $\dim V$*).*

Пусть базис пространства, состоящий из векторов $e_i, \ldots, e_n$, записывается строкой $(e_i, \ldots, e_n)$, а при переходе к матричной записи координаты базисных векторов располагаются в столбец. Матрицей перехода от старого базиса к новому базису $(e_i', \ldots, e_n')$, называется матрица $T = (t_{ij})$, в столбцах которой стоят координаты новых базисных векторов в старом базисе. Таким образом, $(e_i', \ldots, e_n') = (e_i, \ldots, e_n)T$, а координаты вектора x в старом и новом базисах связаны равенствами $x_i = \sum\limits_{j=1}^{n} t_{ij}x_j'$ или, в матричной записи,

$$\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = T \begin{pmatrix} x_1' \\ x_2' \\ \ldots \\ x_n' \end{pmatrix}$$

Задача 12. *1. Пусть x, y — векторы, α, β — скаляры. Доказать, что:*

1) $\alpha x = 0$ тогда и только тогда, когда $\alpha = 0$ или $x = 0$;

2) $\alpha x + \beta y = \beta x + \alpha y$ тогда и только тогда, когда $\alpha = \beta$ или $x = y$.

2. При каких значениях λ:

1) из линейной независимости системы векторов $\{a_i, a_2\}$ вытекает линейная независимость системы $\{\lambda a_i + a_2, a_1 + \lambda a_2\}$;

2) из линейной независимости системы $\{a_i, \ldots, a_n\}$ вытекает линейная независимость системы $\{a_1 + a_2, a_2 + a_3, \ldots, a_{n-1} + a_n, a_n + \lambda a_1\}$.

3. Пусть F — поле, E — его подполе.

1) Доказать, что F является векторным пространством над полем E.

2) Пусть $m_1, \ldots, m_n$ — различные натуральные числа, каждое из которых не делится на квадрат простого числа. Доказать, что числа $1, \sqrt{m_1}, \ldots, \sqrt{m_n}$ линейно независимы в пространстве $\mathbb{R}$ над $\mathbb{Q}$.

3) Пусть $r_1, \ldots, r_n$ — различные рациональные числа из интервала $(0, 1)$. Доказать, что в пространстве $\mathbb{R}$ над полем $\mathbb{Q}$ числа $2^{r_1}, \ldots, 2^{r_n}$ независимы.

3. Пусть векторы $e_i, \ldots, e_n$ и x заданы своими координатами в некотором базисе:

1) $e_i = (l, l, l)$, $e_2 = (1, 1, 2)$, $e_3 = (1, 2, 3)$, $x = (6, 9, 14)$;

2) $e_i = (2, 1, -3)$, $e_2 = (3, 2, -5)$, $e_3 = (1, -1, 1)$, $x = (6, 2, -7)$;

3) $e_i = (1, 2, -1, -2)$, $e_2 = (2, 3, 0, -1)$, $e_3 = (1, 2, 1, 4)$, $e_4 = (1, 3, -1, 0)$, $x = (7, 14, -1, 2)$.

Доказать, что $(e_i, \ldots, e_n)$ — также базис пространства, и найти координаты вектора x в этом базисе.

4. Доказать, что каждая из двух заданных систем векторов S и S' является базисом. Найти матрицу перехода от S к S':

1) $S = ((1,2,1),(2,3,3),(3,8,2))$,
$S' = ((3,5,8),(5,14,13),(1,9,2))$;
2) $S = ((1,1,1,1),(1,2,1,1),(1,1,2,1),(1,3,2,3))$,
$S' = ((1,0,3,3),(-2,-3,-5,-4),(2,2,5,4),(-2,-3,-4,-4))$.

3.1 Ранг матрицы

Назовём пространством столбцов прямоугольной матрицы A размера $m \times n$, линейную оболочку $V_v = (\langle A_1, A_2, \ldots, A_n \rangle$. Размерность $r_v(A) = \dim V$ назовём рангом по столбцам матрицы A. Аналогично вводится ранг по строкам матрицы A $r_h(A) = \dim V_h$, где $V_h = (\langle A^1, A^2, \ldots, A^m \rangle$ — пространство строк матрицы A, т.е. линейная оболочка в $\mathbb{R}^n$, натянутая на векторы-строки.

Теорема 4. *Для любой прямоугольной $m \times n$–матрицы A справедливо равенство $r_h(A) = r_v(A)$ (это число называется рангом матрицы A и обозначается* rank$A$*).*

Доказательство. Доказательство — приведение матрицы к ступенчатому виду. □

Глава 4

Матрицы и линейные отображения

Пусть $\mathbb{R}^n$ и $\mathbb{R}^m$ — векторные пространства столбцов высоты n и m соответственно. Пусть, далее, $A = (a_{ij})$ — матрица размера $m \times n$. Определим отображение $\phi_A : \mathbb{R}^n \to \mathbb{R}^m$, полагая для любого $X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^n$ $\phi_A(X) = x_1A_1 + x_2A_2 + \ldots + x_nA_n$, где $A_1, \ldots, A_n$ — столбцы матрицы A. Так как они имеют высоту m, то в правой части стоит вектор-столбец $Y = [y_1, y_2, \ldots, y_m] \in \mathbb{R}^m$. То есть, $Y = \sum\limits_{j=1}^{n} a_{ij}x_j$, $i = 1, 2, \ldots, m$

Нетрудно проверить, что
1) Если $X = X' + X''$, то $\phi_A(X) = \phi_A(X') + \phi_A(X'')$;
2) $\phi_A(\lambda X) = \lambda\phi_A(X)$.

Определение 20. *Отображение $\phi : \mathbb{R}^n \to \mathbb{R}^m$, обладающее свойствами 1), 2), называется линейным отображением из $\mathbb{R}^n$ в $\mathbb{R}^m$. Часто, в особенности при $n = m$, говорят о линейном преобразовании.*

Предположим теперь, что $\phi : \mathbb{R}^n \to \mathbb{R}^m$ — линейное отображение.

Так как $\mathbb{R}^n = \langle E_1, \ldots, E_n \rangle$ — линейная оболочка стандартных базисных столбцов, имеем

$$X = \sum_{i=1}^{n} x_iE_i.$$

Согласно свойствам 1), 2) имеем

$$\phi(X) = \phi(\sum_{i=1}^{n} x_iE_i) = \sum_{i=1}^{n} x_i\phi(E_i).$$

Последнее соотношение показывает, что отображение ϕ полностью определяется своими значениями на базисных векторах-столбцах. Положив

$\phi(E_j) = [a_{1j}, a_{2j}, \ldots, a_{mj}] = A_j \in \mathbb{R}^m$ мы обнаруживаем, что задание ϕ равносильно заданию прямоугольной матрицы $A = (a_{ij})$ размера $m \times n$ со столбцами $A_1, \ldots, A_m$. То есть, можно положить $\phi = \phi_A$. Матрица A называется матрицей линейного отображения ϕ_A. Суммируем полученные результатаы в утверждении.

Теорема 5. *Между линейными отображениями $\mathbb{R}^n$ в $\mathbb{R}^m$ и матрицами размера $m \times n$ существует взаимно однозначное соответствие.*

4.1 Линейная комбинация матриц

Линейные функции, равно как и произвольные линейные отображения $\mathbb{R}^n \to \mathbb{R}^m$ при фиксированных m и n можно складывать и умножать на скаляры. В самом деле, пусть $\phi_A, \phi_B : \mathbb{R}^n \to \mathbb{R}^m$ — два линейных отображения. Тогда отображение $\phi = \alpha\phi_A + \beta\phi_B : \mathbb{R}^n \to \mathbb{R}^m$, $\alpha, \beta \in \mathbb{R}$ определено своими значениями $\phi(X) := \alpha\phi_A(X) + \beta\phi_B(X)$. Очевидно, ϕ — линейное отображение. Следовательно, определена матрица C этого отображения. Столбец C_j можно определить из соотношения $C_j = [c_{1j}, c_{2j,\ldots,c_{mj}}] = \phi(E_j) = \alpha\phi_A(E_j) + \beta\phi_B(E_j) = \alpha A_j + \beta B_j$. Следовательно, $c_{ij} = \alpha a_{ij} + \beta b_{ij}$.

4.2 Умножение матриц

Пусть $\phi_B : \mathbb{R}^n \to \mathbb{R}^s$, $\phi_A : \mathbb{R}^s \to \mathbb{R}^m$ — два линейных отображения с матрицами $A = (a_{ij})_{i=1,\ldots m, j=1,\ldots,s}$, $B = (b_{ij})_{i=1,\ldots s, j=1,\ldots,n}$. Рассмотрим $\phi_C = \phi_A \circ \phi_B : \mathbb{R}^n \to \mathbb{R}^m$. Пусть $X = [x_1, \ldots, x_n] \in \mathbb{R}^n$, $Y = [y_1, \ldots, y_s] \in \mathbb{R}^s$ и $Z = [z_1, \ldots, z_m] \in \mathbb{R}^m$. При этом $z_i = \sum\limits_{k=1}^{s} a_{ik} y_k = \sum\limits_{k=1}^{s} a_{ik} \sum\limits_{j=1}^{n} b_{kj} x_j = \sum\limits_{j=1}^{n} (\sum\limits_{k=1}^{s} a_{ik} b_{kj}) x_j$. С другой стороны, $z_i = \sum\limits_{j=1}^{n} c_{ij} x_j$. То есть, $c_{ij} = \sum\limits_{k=1}^{s} a_{ik} b_{kj}$

Теорема 6. *Произведение $\phi_A\phi_B$ двух линейных отображений с матрицами A и B является линейным отображением с матрицей $C = AB$. Другими словами, $\phi_A\phi_B = \phi_{AB}$.*

Следствие 1. *Умножение матриц ассоциативно:*

доказательство — упражнение.

4.3 Транспонирование матриц

Пусть $A = (a_{ij})_{i=1,\dots m, j=1,\dots,n}$. Тогда транспонированной матрицей ${}^tA$ называется матрица с компонентами $a^t_{ij} = a_{ji}$, $i = 1, \dots, n$, $j = 1, \dots m$.

Основное свойство транспонирования — ${}^t(AB) = {}^t B\, {}^t A$.

Доказательство — упражнение.

4.4 Ранг прозведения матриц

Теорема 7. $\operatorname{rank}(AB) \le \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$

Доказательство. Для строк и столбцов матрицы C имеем соотношения $C_i = AB_i$ и $C^j = A^j B$. Тогда первое соотношение влечет, что линейная зависимость системы B_i влечет линейную зависимость C_i, следовательно, $\operatorname{rank}(C) = \operatorname{rank}(C_i) \le \operatorname{rank}(B_i) = \operatorname{rank}(B)$. Аналогично, второе соотношение влечет $\operatorname{rank}(C) \le \operatorname{rank}(A)$. □

Матричная единица E_{ij} — матрица, у которой на i, j месте стоит 1, остальные элементы — нули.

Символ Кронекера $\delta_{ij} = \begin{pmatrix} 1 & i = j \\ 0 & i \neq j \end{pmatrix}$

Задача 13. *1. Перемножить матрицы а)*

$$\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$$

b)

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix}$$

c)

$$\begin{pmatrix} 3 & -4 & 5 \\ 2 & -3 & 1 \\ 3 & -5 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 29 \\ 2 & 18 \\ 0 & -3 \end{pmatrix}$$

d)

$$\begin{pmatrix} 1 & 5 & 3 \\ 2 & -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & -3 & 5 \\ -1 & 4 & -2 \\ 3 & -1 & 1 \end{pmatrix}$$

2. Вычислить а)

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}^n$$

b)

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^n$$

3. Вычислить степени квадратной матрицы

$$H = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

4. Вычислить $e^A = Id + \frac{A}{1!} + \frac{A^2}{2!} + \dots$ *для a)* $A = \begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix}$

b) $A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{pmatrix}$

5. a) Доказать, что $E_{ij}E_{pq} = \delta_{jp}E_{iq}$

b) Пусть A *— произвольная матрица. Вычислить* $E_{ij}A$.

c) Пусть A *— произвольная матрица. Вычислить* AE_{ij}.

d) Пусть A *— квадратная матрица, причём* $E_{ij}A = AE_{ij}$ *для всех матричных единиц* E_{ij}. *Доказать, что* $A = \lambda Id$ *для некоторого скаляра* λ.

e) Пусть A *— квадратная матрица, причём* $E_{ii}A = AE_{ii}$ *для всех* i. *Доказать, что матрица* A *диагональна.*

f) Найти все матрицы A *порядка* n *такие, что* $\mathrm{tr}AX = 0$ *для любой матрицы* X *порядка* n. *Здесь* $\mathrm{tr}A = \sum_{i=1}^{n} a_{ii}$.

Глава 5

Квадратные матрицы

Множество всех квадратных матриц (a_{ij}) порядка n с вещественными коэффициентами, обычно обозначается $M_n(\mathbb{R})$ (или M_n). Можно показать, что M_n — векторное пространство. При этом по отношению к операциям сложения и умножения матриц выполнены свойства ассоциативности и дистрибутивности.

Определение 21. *Говорят, что квадратные матрицы фиксированного порядка n образуют матричное (ассоциативное) кольцо. Кроме того, с учётом легко проверяемых правил $\lambda AB = (\lambda A)B = A(\lambda B)$ умножения на скаляры $\lambda \in \mathbb{R}$ множество M_n называют также алгеброй матриц над $\mathbb{R}$.*

Рассмотрим единичную матрицу $E = \mathrm{Id} = \delta_{kj}$, где

$$\delta_{kj} = \begin{cases} 1 & k = j \\ 0 & k \neq j. \end{cases}$$

— символ Кронекера. Очевидно, что $\mathrm{rank}(E) = n$.

Правило умножения матриц, в котором следует заменить b_{kj} на δ_{kj}, показывает, что справедливы соотношения $\forall A \in M_n$, $EA = A = AE$.

Для данной матрицы $A \in M_n(\mathbb{R}$ можно попробовать найти такую матрицу $A' \in M_n(\mathbb{R})$, чтобы выполнялись соотношения $AA' = E = A'A$. Сразу же заметим, что $AA' = E = A''A \Rightarrow A'' = A'$.

Действительно, $A'' = A''E = A''(AA') = (A''A)A' = EA' = A'$. Таким образом, матрица A', если она существует, единственна. Её называют матрицей, обратной к A, и обозначают A^{-1}:

$$AA^{-1} = E = A^{-1}A \tag{5.1}$$

При выполнении (5.1) говорят ещё, что матрица A обратима. Определение.

Определение 22. *Матрица $A \in M_n(\mathbb{R})$ называется невырожденной, если система её строк (а тем самым и столбцов) линейно независима, т.е.* $\mathrm{rank}A = n$. *Если* $\mathrm{rank}A < n$, *то A называется вырожденной.*

Теорема 8. *Матрица $A \in M_n(\mathbb{R})$ обратима тогда и только тогда, когда она невырожденна.*

Доказательство. 1) ($\Rightarrow$) Если $AB = E$ (или $BA = E$), то по теореме 7 имеем $n = \mathrm{rank}E = \mathrm{rank}AB \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\} \leq n$, откуда $\mathrm{rank}A = n$.

2) ($\Leftarrow$) Если $\mathrm{rank}A = n$, то $\langle E_1, \ldots, E_n\rangle = \mathbb{R}^n = \langle a_1, \ldots, A_n\rangle$, следовательно, $E_j = \sum\limits_{i=1}^{n} a'_{ij} A_i$, $j = 1, \ldots, n$. Тогда $E = AA'$. □

5.1 Алгоритм поиска обратной матрицы

Рассмотрим в $M_n(\mathbb{R})$ так называемые элементарные матрицы следующих типов:

$$F_{s,t} = \mathrm{Id} - E_{ss} - E_{tt} + E_{st} + E_{ts};$$

$$F_{s,t}(\lambda) = \mathrm{Id} + \lambda E_{st};$$

$$F_s(\lambda) = \mathrm{Id} + \lambda E_{st} = \mathrm{diag}\{1, \ldots, \lambda, 1, \ldots, 1\}, \lambda \neq 0.$$

Пусть A — произвольная $m \times n$-матрица. Тогда непосредственно проверяется, что матрица $A' = FA$ получается из A посредством элементарного преобразования (э.п.) над строками типа (I) или (II) в зависимости от того, будет $F = F_{st}$ или $F = F_{st}(\lambda)$. В случае $F = F_s(\lambda)$ будем говорить об э.п. типа (III) (умножение s-й строки A на λ). Аналогично, матрица $A'' = AF$ получается из A посредством э.п. столбцов.

Известно, что э.п. типов (I) и (II), совершаемыми над строками и столбцами, A приводится к матрице с диагональной невырожденной подматрицей. Поскольку

$$\begin{pmatrix} a_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ldots & \ldots & a_r & 0 & \ldots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix} = F_1(a_1) \cdots F_r(a_r) \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ldots & \ldots & 1 & 0 & \ldots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix},$$

использование э.п. типа (III) даёт возможность получить из A матрицу вида $\begin{pmatrix} \mathrm{Id}_r & 0 \\ 0 & 0 \end{pmatrix}$ (здесь Id_r — единичная матрица в $M_r(\mathbb{R})$. Таким образом, $P_k P_{k-1} \cdots P_1 A Q_1 Q_2 \cdots Q_l = \begin{pmatrix} \mathrm{Id}_r & 0 \\ 0 & 0 \end{pmatrix}$, где P_i, (соответственно Q_j) – элементарные матрицы порядка m (соответственно n).

Предложение 8. *Всякая невырожденная $n \times n$-матрица записывается в виде произведения элементарных матриц.*

Доказательство. Действительно, все невырожденные матрицы порядка n элементарными преобразованиями приводятся к диагональному виду, поскольку их ранги равны n. Соотношение $P_k P_{k-1} \cdots P_1 A Q_1 Q_2 \cdots Q_l = \mathrm{Id}$, переписанное в виде $A = P_1^{-1} \cdots P_k^{-1} Q_l^{-1} \cdots Q_1^{-1}$даёт нужное утверждение. □

Если в приведенных выше рассуждениях ограничиться преобразованиями над строками и рассмотреть с самого начала расширенную матрицу $(A|\mathrm{Id})$ размера $n \times 2n$, то в случае невырожденной матрицы $M_n(\mathbb{R})$ возникнет цепочка $(A|\mathrm{Id}) \xrightarrow{P_1} (P_1 A|P_1 \mathrm{Id}) \xrightarrow{P_2} (P_2 P_1 A|P_2 P_1 \mathrm{Id}) \xrightarrow{P_3} \ldots \xrightarrow{P_k} (P_k \cdots P_2 P_1 A|P_k \cdots P_2 P_1 \mathrm{Id}) = (\mathrm{Id}|A^{-1})$. Она оборвётся на k-м шаге, когда в левой половине расширенной матрицы место A заполнит единичная матрица Id. В правой половине при этом получится однозначный ответ: A^{-1}. В случае вырожденной матрицы A процесс оборвётся, возможно, раньше — приведением A к ступенчатому виду и вычислением ранга $r = \mathrm{rank} A$.

Пример 9. *Пусть*

$$A - \begin{pmatrix} 0 & 2 & 0 \\ 1 & 1 & -1 \\ 2 & 1 & -1 \end{pmatrix}$$

Имеем

$$(A|\mathrm{Id} - \left(\begin{array}{ccc|ccc} 0 & 2 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 & 0 \\ 2 & 1 & -1 & 0 & 0 & 1 \end{array}\right)$$

Глава 6

Группа перестановок

Пусть Ω — конечное множество из n элементов. Поскольку природа его элементов для нас несущественна, удобно считать, что $\Omega = \{1, 2, \ldots, n\}$. Элементы множества $S_n = S(\Omega)$ всех взаимно однозначных преобразований $\Omega \to \Omega$, называются перестановками. Произвоную перестановку $\pi \in S_n$ можно представить в виде

$$\pi = \begin{pmatrix} 1 & 2 & \ldots & n \\ \pi(1) & \pi(2) & \ldots & \pi(n) \end{pmatrix}$$

Умножение перестановок. $(\sigma\tau)(i) = \sigma(\tau(i))$, $i = 1, \ldots, n$. Наделенное такой операцией, множество S_n становится группой (проверить самосоятельно).

Определение 23. *Цикл перестановки σ — элементарная составляюая перестановки, описывающая переход неторого элемента Ω при последовательном применении перестаовки σ.*

Определение 24. *Цикл длины 2 наывается транспозицией.*

Предложение 9. *Каждая перестановка является произведением транспозиций.*

6.1 Знак перестановки

Теорема 9. *Пусть π — перестановка из S_n и $\pi = \tau_1\tau_2\ldots\tau_k$ — произвольное разложение π в произведение транспозиций. Тогда число* $\mathrm{sgn}(\pi) := (-1)^k$*, называемое знаком π (иначе: сигнатурой или чётностью), полностью определяется перестановкой π и не зависит от способа разложения, т.е. чётность целого числа k для данной перестановки π всегда одна и та же. Кроме того,* $\mathrm{sgn}(\alpha\beta) = \mathrm{sgn}(\alpha)\mathrm{sgn}(\beta)$.

Доказательство. 1) Предположим, что наряду с данным мы имеем также разложение $\pi = \tau_1^{'}\tau_2^{'}\dots\tau_{k'}^{'}$, причём четности k и k' различны. Это значит, что целое число $k + k'$ нечётно. Так как $(\tau_s^{'})^2 = e$, то, последовательно умножая справа обе части равенства $\tau_1\tau_2\dots\tau_k = \tau_1^{'}\tau_2^{'}\dots\tau_{k'}^{'}$, на $\tau_{k'}^{'},\dots\tau_2^{'},\tau_1^{'}$, получим $\tau_1\tau_2\dots\tau_k\tau_{k'}^{'},\dots\tau_2^{'},\tau_1^{'} = e$.

То есть задача сведена к следующей. Пусть $e = \sigma_1\sigma_2\dots\sigma_m$ — запись единичной перестановки в виде произведения $m > 0$ транспозиций. Нужно показать, что обязательно m — чётное число. С этой целью будет установлено, что от данной записи мы можем перейти к записи e в виде произведения $m - 2$ транспозиций. Продолжив этот спуск, мы пришли бы при нечётном m к одной транспозиции τ. Но, очевидно, $e \neq \tau$. Итак, нам нужно обосновать спуск от m к $m - 2$ множителям.

2) Пусть s, $1 \leq s \leq n$ — любое фиксированное натуральное число, входящее в одну из транспозиций $\sigma_2,\dots,\sigma_m$. Для определённости считаем, что $e = \sigma_1\dots\sigma_{p-1}\sigma_p\sigma_{p+1}\dots\sigma_m$, где $\sigma_p = (st)$, а $\sigma_{p+1},\dots,\sigma_m$ не содержат s. Для σ_{p-i} имеются четыре возможности:

а) $\sigma_{p-i} = (st)$; тогда отрезок $\sigma_{p-1}\sigma_p$ из записи e удаляется, и мы приходим к $m - 2$ транспозициям;

б) $\sigma_{p-i} = (sr)$, $r \neq s,t$, здесь $\sigma_{p-1}\sigma_p = (sr)(st) = (st)(rt)$, и мы сдвинули вхождение s на одну позицию влево, не изменив m;

в) $\sigma_{p-i} = (tr)$,$r \neq s,t$, здесь $\sigma_{p-1}\sigma_p = (tr)(st) = (sr)(tr)$, и снова, как в случае б), произошёл сдвиг s влево без изменения m;

г) $\sigma_{p-i} = (qr)$, $\{q,r\}\bigcap\{s,t\} = \emptyset$; здесь $\sigma_{p-1}\sigma_p = (qr)(st) = (st)(qr)$.

В случае а) наша цель достигнута. В случаях б)-г) повторяем процесс, сдвигая вхождение s на одну позицию влево. В конечном счёте мы придем либо к случаю а), либо к экстремальному случаю, когда $e = \sigma_1^{'}\sigma_2^{'}\dots\sigma_m^{'}$, причём $\sigma_1^{'} = (st')$ и s не имеет вхождений в $\sigma_2^{'},\dots,\sigma_m^{'}$. Значит, $\sigma_k^{'}(s) = s$ при $k > 1$ и $s = e(s) = \sigma_1^{'}(s) = t' \neq s$. Полученное противоречие доказывает утверждение об инвариантности $\mathrm{sgn}(\pi)$.

3) Если $\alpha = \tau_1\dots\tau_k$, $\beta = \tau_{k+1}\dots\tau_{k+l}$ то $\alpha\beta = \tau_1\dots\tau_k\tau_{k+1}\dots\tau_{k+l}$ и $\mathrm{sgn}(\alpha) = (-1)^k$, $\mathrm{sgn}(\beta) = (-1)^l$ $\mathrm{sgn}(\alpha\beta) = (-1)^{k+l} = \mathrm{sgn}(\alpha)\mathrm{sgn}(\beta)$. □

Определение 25. *Перестановка $\pi \in S_n$ называется чётной, если* $\mathrm{sgn}(\pi) = 1$, *и нечётной, если* $\mathrm{sgn}(\pi) = -1$.

Из определения вытекает, что все транспозиции — нечётные перестановки, а $\mathrm{sgn}(e) = 1$.

Глава 7

Определители

Если A — квадратная таблица, заполненная своими коэффициентами (обычно числами), то определитель порядка n — это число (или выражение), приписываемое матрице A и определённое формулой полного развёртывания

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)} \tag{7.1}$$

Другими словами, определителем $\det A$ матрицы $A = (a_{ij})_{i,j=1}^{n}$ называется алгебраическая сумма всевозможных произведений коэффициентов a_{ij}, взятых по одному из каждой строки и из каждого столбца. В каждом произведении сомножители записываются в порядке следования строк, а номера столбцов определяются образами $\sigma(1), \sigma(2), \dots, \sigma(n)$ номеров строк при перестановке $\sigma \in S_n$. Всего под знаком суммы в (7.1) стоит $n!$ слагаемых; слагаемые, отвечающие чётным перестановкам, входят со знаком плюс, а отвечающие нечётным перестановкам, — со знаком минус.

7.1 Свойства функции det

Теорема 10. *Определители любой квадратной матрицы A и транспонированной с ней матрицы A^t совпадают:* $\det A^t = \det A$.

Доказательство. Положив $A = (a_{ij})$, ${}^tA = (a'_{ij})$, где $a'_{ij} = a_{ij}$, и заметив, что $k = \pi(\pi^{-1}(k))$ для любой перестановки $\pi \in S_n$ и для любого номера $k \in \{1, 2, \dots, n\}$, мы видим, что упорядочение множителей произведения $a'_{1\pi(1)} \dots a'_{n\pi(n)}$ в соответствии с перестановкой π^{-1} даёт

$$a'_{1\pi(1)} \dots a'_{n\pi(n)} = a'_{\pi^{-1}(1)\pi(\pi^{-1}(1))} \dots a'_{\pi^{-1}(n)\pi(\pi^{-1}(n))} = a'_{\pi^{-1}(1)1} \dots a'_{\pi^{-1}(n)n} = a_{1\pi^{-1}(1)} \dots a_{n\pi^{-1}(n)}.$$

Так как $\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1})$ ($\operatorname{sgn}(\pi)\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi\pi^{-1}) = \operatorname{sgn}(e) = 1$) и $\{\pi^{-1} | \pi \in S_n\} = S_n$, по формуле (7.1) имеем

$$\overset{t}{\det} A = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a'_{1\pi(1)} \dots a'_{n\pi(n)} = \sum_{\pi \in S_n} \operatorname{sgn}(\pi^{-1}) a'_{1\pi^{-1}(1)} \dots a'_{n\pi^{-1}(n)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \dots a$$

□

Теорема 11. *Функция* $\det : A \mapsto \det A$ *на множестве* $M_n(\mathbb{R})$ *обладает следующими свойствами.*

l. $\det A$ *— кососимметрическая функция строк матрицы* A *(т.е. при перестановке местами любых двух строк определитель меняет знак на противоположный).*

2. $\det A$ *— полилинейная функция строк матрицы* A *(т.е. определитель матрицы* A *является линейной функцией элементов любой её строки* A^i*.*

3. $\det E = 1$*.*

Доказательство. 1. Пусть A' — матрица, получающаяся из A перестановкой строк A^s и A^t, т.е. $A'^s = A^t$, $A'^t = A^s$, $A'^i = A^i$ при $i \neq s, t$. Тогда, записав любую перестановку $\pi \in S_n$ в виде $\pi = \sigma\tau$ с транспозицией $\tau = (s, t)$, будем иметь

$$\det A' = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a'_{1\pi(1)} \dots a'_{n\pi(n)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma\tau) a'_{1\sigma\tau(1)} \dots a'_{s\sigma\tau(s)} \dots a'_{t\sigma\tau(t)} \dots a'_{n\sigma\tau(n)} = \sum_{\sigma \in S_n} s$$

2. Пусть $A = (a_{ij})$, и пусть $A^k = \lambda' A'^k + \lambda'' A''^k$, где штрихи указывают на вспомогательные матрицы

$$A' = [A^1, \dots, A^{k-1}, A^{k'}, A^{k+1}, \dots, A^n],$$

$$A'' = [A^1, \dots, A^{k-1}, A^{k''}, A^{k+1}, \dots, A^n].$$

По условию $a_{kj} = \lambda' a'_{kj} + \lambda'' a''_{kj}$, $j = 1, 2, \dots, n$. По определению $\det[A^1, \dots, A^k, \dots, A^n] = \det A = \sum\limits_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)} = \sum\limits_{\sigma \in S_n} p_\sigma a_{k,\sigma(k)}$, где $p_\sigma \sigma \in S_n$, — коэффициенты, не зависящие от элементов строки A^k. Собирая подобные члены, отвечающие тем $\sigma \in S_n$, для которых $\sigma(k) = j$, и полагая $\alpha_j = \sum\limits_{\sigma(k)=j} p_\sigma$, получим нужное свойство линейности:

$$\det[\dots, A^k, \dots] = \sum_{j=1}^{n} \alpha_j a_{kj},$$

$$\det[\ldots, A'^k+\lambda''A''^k, \ldots] = \sum_{j=1}^{n} \alpha_j(\lambda' a'_{kj}+\lambda'' a''_{kj}) = \lambda' \sum_{j=1}^{n} \alpha_j a'_{kj}+\lambda'' \sum_{j=1}^{n} \alpha_j a''_{kj} = \lambda' \det(A')+\lambda'' \det(A'').$$

3. Очевидно, $\det E = \sum\limits_{\sigma\in S_n} \operatorname{sgn}(\sigma)\delta_{1\sigma(1)}\ldots\delta_{n\sigma(n)} = \operatorname{sgn}(e)\delta_{11}\ldots\delta_{nn}$. □

Некоторые дополнительные свойства есть следствия уже известных.

Следствие 2. *4. Пусть $A \in M_n(\mathbb{R})$, $\lambda \in \mathbb{R}$. Тогда*

$$\det(\lambda A) = \lambda^n \det(A)$$

5. Определитель с нулевой строкой равен нулю.

6. Если в квадратной матрице A две строки совпадают, то её определитель равен нулю.

7. Определитель не меняется, если над его строками совершать элементарные преобразования следующего типа: $A'^i = A^i$ для всех $i \neq s$ и $A'^s = A^s + \lambda A^t$, $s \neq t$, $\lambda \in \mathbb{R}$.

Доказательство. 4. Действительно, в силу свойства 2, применённого последовательно к строкам с номерами $1, 2, \ldots, n$, имеем $\det(\lambda A) = \det[\lambda A^1, \ldots, \lambda A^n] = \lambda \det[A^1, \lambda A^2, \ldots, \lambda A^n] = \ldots = \lambda^n \det(A)$.

5. Пусть, например, $A^k = (0, \ldots, 0$. Тогда и $2A^k = (0, \ldots, 0)$. Следовательно, по 2. $\det(A) = 2\det(A)$, то есть $\det(A) = 0$.

6. Поменяв местами две совпадающие строки A^i, A^j в A, мы получим ту же матрицу A. С другой стороны, согласно свойству 1 для det значение $\det(A)$ примет противоположный знак. Таким образом, $\det(A) = -\det(A)$, откуда $2\det(A) = 0$ и $\det(A) = 0$.

7. Достаточно рассмотреть случай применения одного элементарного преобразования. Пусть после прибавления к s-й строке матрицы A её t-й строки, умноженной на λ, получилась матрица A'. Тогда в соответствии со свойствами 1 и 6 для det имеем

$$\det(A') = \det[A^1, \ldots, A^s+\lambda A^t, \ldots, A^n] = \det[A^1, \ldots, A^s, \ldots, A^n]+\det[A^1, \ldots, \lambda A^t, \ldots, A^n] = \det(A)-$$

□

7.2 Разложение определителя по элементам столбца или строки

Существует регулярный способ вычисления определителей, основанный на редукции к определителям меньшего порядка. При этом используются понятия минора и алгебраического дополнения.

Определение 26. *Определитель матрицы, получающейся из $A = (a_{st})$ вычёркиванием i-й строки и j-го столбца, обозначается M_{ij} и называется минором матрицы A, соответствующим элементу a_{ij}. Величина $A_{ij} = (-1)^{i+j} M_{ij}$ называется алгебраическим дополнением элемента a_{ij}.*

Предложение 10. *Если*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

то $\det(A) = a_{11} M_{11} = a_{11} A_{11}$.

Доказательство. По теореме 10 $\det A = \det^t A$ и $a_{\pi(1),1} = 0$ для любой перестановки $\pi \in S_n$, для клтоторой $\pi(!) \neq 1$, имеем

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{\pi(1),1} \dots a_{\pi(n),n} = \sum_{\pi \in S_n, \pi(1)=1} \operatorname{sgn}(\pi) a_{11} a_{\pi(2),2} \dots a_{\pi(n),n}.$$

Совокупность всех перестановок $\pi \in S_n$, оставляющих на месте символ 1, отождествляется с множеством S_{n-i} перестановок, действующих на множестве $\{2, 3, \dots, n\}$. Таким образом,

$$\det(A) = a_{11} \sum_{\pi \in S_{n-1}} \operatorname{sgn}(\pi) a_{\pi(2),2} \dots a_{\pi(n),n} = a_{11} M_{11}$$

□

Теорема 12. *Пусть $A = (a_{ij}) \in M_n(\mathbb{R})$. Справедливы следующие формулы:*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij}$$ *(разложение определителя по элементам j-го сто*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij} = \sum_{j=1}^{n} a_{ij} A_{ij}$$ *(разложение определителя по элементам i-й стр*

Иначе говоря, определитель матрицы A равен сумме произведедений всех элементов некоторого столбца (некоторой строки) на их алгебраические дополнения.

Доказательство. 1) Опираясь на основные свойства 1 и 2 определителей (сначала относительно столбцов, а затем относительно строк), выпишем цепочку равенств:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1j} & a_{1n} \\ a_{12} & \dots & 0 & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & 0 & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \dots & 0 & a_{1n} \\ a_{12} & \dots & a_{2j} & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & 0 & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \dots & 0 & a_{1n} \\ a_{12} & \dots & 0 & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nj} & a_{nn} \end{vmatrix}$$

Последнее равенство основано на Предложении 10.

2) Вторая формула получается из первой с использованием Теоремы 10. □

Глава 8

Жорданова нормальная форма матрицы

8.1 Проекторы

Пусть $V = W_1 \oplus W_2 \oplus \ldots \oplus W_m$ — разложение пространства V в прямую сумму m подпространств. Тогда каждый вектор $x \in V$ однозначно записывается в виде $x = x_i + x_2 + \ldots + x_m$, $x_i \in W_i$, а отображение $P_i : x \to x_i$ является линейным оператором на V. Кроме того,

$$P_1 + \ldots + P_m = Id,$$

причём $P_iP_j = 0$ при $i \neq j$ и $P_i^2 = P_i$. Наконец, $W_i = P_iV = \{x \in V | P_ix = x\}$, $K_i = \mathrm{Ker}P_i = W_1 + \ldots + \hat{W}_i + \ldots + W_m$ и P_i суть оператор проектирования V на W_i вдоль K_i.

Теорема 13. *Пусть $P_1, \ldots, P_m : V \to V$ — конечное множество линейных операторов, удовлетворяющих условиям*

$$\sum_{i=1}^{m} P_i = Id; P_i^2 = P_i, 1 \leq i \leq m; P_iP_j = 0, i \neq j. \tag{8.1}$$

Тогда $V = W_1 \oplus W_2 \oplus \ldots \oplus W_m$, где $W_i = \mathrm{Im}P_i$.

Доказательство. По условию для любого $x \in V$ имеем $x = \mathrm{Id}x = \sum\limits_{i=1}^{m} P_ix = x_i + \ldots + x_m$, $x_i \in W_i$. Поэтому $V = W_1 + \ldots + W_m$.

Эта сумма является прямой. Именно, предположим, что $x \in W_j \bigcap \sum\limits_{i=1, i\neq j}^{m} W_i$. Так как $W_i = \mathrm{Im}P_i$, то найдутся такие векторы $x_1, \ldots, x_m$, что

$$x = P_j(x_j) = \sum_{i=1, i\neq j}^{m} P_i(x_i).$$

Применяя к этому равенству оператор P_j и используя определяющие свойства $P_j^2 = P_j$, $P_jP_i = 0$ при $i \neq j$, получим $x = P_j(x_j) = P_j^2(x_j) = \sum\limits_{i=1,i\neq j}^{m} P_jP_i(x_j) = 0$. Таким образом, сумма $V = \sum\limits_{i=1}^{m} W_i$ прямая и P_i — оператор проектирования V на W_i вдоль $K_i = \mathrm{Ker}P_i = \sum\limits_{i=1,i\neq j}^{m} W_j$. $\square$

Добавим, что если $P^2 = P$ и $V = U \oplus W$ — связанное с этим проектором прямое разложение с $U = \mathrm{Im}P = \langle e_1, \ldots, e_r\rangle$, $W = \mathrm{Ker}P = \langle e_{r+i}, \ldots, e_n\rangle$, то в выбранном базисе оператору P отвечает матрица

$$\begin{pmatrix} Id_r & 0 \\ 0 & 0 \end{pmatrix}, r = \mathrm{rank}P \tag{8.2}$$

В частности, мы видим, что любая $n \times n$-матрица A ранга r, обладающая свойством $A^2 = A$, подобна матрице P: $B^{-1}aB = P$ и $\mathrm{rank}A = \mathrm{tr}A$.

8.2 Инвариантные подпространства

Всякий линейный оператор $A : V \to V$ действует не только на отдельные векторы $x \in V$, но и на подпространства $U \subset V$: $AU = \{Ax | x \in U\}$. В связи с этим важное значение приобретает понятие инвариантности.

Определение 27. *Подпространство $U \subset V$ инвариантно относительно линейного оператора $A : V \to V$, если $AU \subset U$.*

Наличие собственного инвариантного подпространства $U \subset V$ даёт возможность упростить матрицу A оператора L_A путём выбора надлежащего базиса в V. Именно, если дополнить базис $(e_1, \ldots, e_r)$ в U до базиса $(e_1, \ldots, e_r, e_{r+1}, \ldots, e_n)$ в Y, то из условия $Ae_i \in U$, $1 \leq i \leq r$, следует, что в этом базисе матрицей оператора L_A будет

$$A = \begin{pmatrix} A_1 & A_0 \\ 0 & A_2 \end{pmatrix},$$

где A_1 — $r \times r$-матрица, A_2 — $(n-r)x(n-r)$-матрица и A_0 — $r \times (n-r)$-матрица. На A_1 можно смотреть как на матрицу линейного оператора L_{A_U} — оператора L_A, ограниченного на U (удобно положить $A_1 = A_U$). Пусть A_0 — нулевая матрица. Тогда, очевидно, $W = (e_{r+1}, \ldots, e_n)$ тоже будет инвариантным подпространством в V, а A_2 — матрицей оператора L_{A_W} В этом случае говорят о прямой сумме операторов, соответствующей разложению $V = U \oplus W$ в прямую сумму инвариантных

подпространств. Матрица прямой суммы операторов имеет клеточно-диагональный вид:

$$A = \begin{pmatrix} A_U & 0 \\ 0 & A_W \end{pmatrix} \tag{8.3}$$

Таким образом, доказана

Теорема 14. *Пространство V является прямой суммой двух подпространств U, W, инвариантных относительно линейного оператора $A : V \to V$, тогда и только тогда, когда матрица этого оператора в какомлибо базисе принимает клеточно-диагональный вид (8.3)*

Задача 14. *1. Найти собственные векторы и собственные значения:*

а) оператора дифференцирования в пространстве $\mathbb{R}[X]_n$;

б) оператора $X \to X^{tr}$ в пространстве $Mat_n(\mathbb{R})$;

в) оператора $x\frac{d}{dx}$ в пространстве $\mathbb{R}[X]_n$;

г) оператора $\frac{1}{x}\int\limits_0^x f(t)dt$ в пространстве $\mathbb{R}[X]_n$;

8.3 Собственные векторы

Определение 28. *Любой ненулевой вектор из одномерного подпространства, инвариантного относительно A, называется собственным вектором оператора A. Если x — собственный вектор: $Ax = \lambda x$, то скаляр $\lambda \in \mathbb{R}$ называется собственным значением оператора A, отвечающим собственному вектору x. Иногда говорят также: характеристический вектор, характеристическое значение.*

Заметим, что $Ax = \lambda x \Rightarrow A^k x = \lambda^k x$, откуда $f(Ax) = f(\lambda)x$ каков бы ни был многочлен f. В частности, $f(A) = 0 \Rightarrow f(\lambda) = 0$ для всякого собственного значения λ оператора A. Пусть $V^\lambda = \{v \in V | Av = \lambda v\}$ — подпространство, состоящее из 0 и всех собственных векторов, ассоциированных с собственным значением λ.

Определение 29. *Очевидная импликация $Ax = \lambda x$, $Ay = \lambda y \Rightarrow \forall \alpha, \beta \in \mathbb{R}$ $A(\alpha x + \beta y) = \lambda(\alpha x + \beta y)$ даёт основание называть V^λ собственным подпространством оператора A, ассоциированным с λ. Его размерность $\dim V^\lambda$ называется геометрической кратностью собственного значения λ.*

Условие существования собственного вектора записывается, очевидно, в виде

$$(A - \lambda Id)x = 0, x \neq 0, \tag{8.4}$$

т.е. $\mathrm{Ker}(A - \lambda Id) \neq 0$.

Это значит, что оператор $A - \lambda Id$ вырожден:

$$\det(A - \lambda Id) = 0. \tag{8.5}$$

Если в каком-нибудь базисе (e_i) пространства V матрицей оператора L_A является $A = (a_{ij})$, то матрицей оператора $L_{A-\lambda Id}$ будет $A - \lambda Id$, так что условие (8.5) переписывается в виде

$$\det(A - \lambda Id) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0$$

Расписав определитель, получим многочлен

$$\chi_A(t) = \det(tId - A) = t^n + \chi_1 t^{n-1} + \ldots + \chi_{n-1} t + \chi_n \tag{8.6}$$

степени n относительно независимой переменной t с коэффициентами $\chi_i \in \mathbb{R}$.

Определение 30. *Многочлен (8.6) называется характеристическим многочленом матрицы A. Уравнение $\chi_A(t) = 0$ называется также характеристическим.*

Теорема 15. *Характеристические многочлены подобных матриц совпадают.*

Доказательство. Пусть $A' = C^{-1}AC$. Тогда $\det(tId - A') = \det(tC^{-1}IdC - C^{-1}AC) = (C^{-1}(tId - A)C) = \det C^{-1} \det(tId - A) \det C = \det(tId - A)$. $\square$

Итак, полагаем $\chi_{L_A}(t) := \chi_A(t)$.

Определяющее равенство (8.6) показывает, что скаляр $\lambda \in F$ является собственным значением оператора L_A тогда и только тогда, когда $\chi_A(\lambda) = 0$, т.е. λ — корень характеристического многочлена. Если многочлен $\chi_A(t)$ не имеет корней в F, то у оператора L_A нет собственных векторов. Всякий линейный оператор, действующий на комплексном векторном пространстве, обладает собственными векторами.

Определение 31. *Кратность λ как корня характеристического многочлена $\chi_A(t)$ называется алгебраической кратностью собственного значения λ оператора L_A.*

Теорема 16. *Геометрическая кратность собственного значения λ не превосходит его алгебраической кратности.*

Доказательство. По определению геометрическая кратность есть размерность m пространства V^λ решений уравнения $L_A x = \lambda x$. Очевидно, что V^λ инвариантно относительно L_A, и если $L_A^{'}$ — ограничение L_A на V^λ, то $\det(tId' - A1) = (t-\lambda)^m$, причём $\chi_{L_A}(t) = (t-\lambda)^m q(t)$, где $q(t)$ — некоторый многочлен из $F[t]$. Пусть λ — корень кратности $k \geq 0$ многочлена $q(t)$. В таком случае алгебраической кратностью λ будет $m+k$. $\square$

Задача 15. *1. Доказать, что если оператор A^2 имеет собственное значение λ^2, то одно из чисел λ и $-\lambda$ является собственным значением оператора A.*

2. Доказать, что если A и B — квадратные матрицы одинаковых порядков, то матрицы AB и BA имеют совпадающие характеристические многочлены.

3. Доказать, что все характеристические числа матрицы отличны от нуля тогда и только тогда, когда матрица невырожденная.

4. Найти собственные значения и собственные векторы линейных операторов, заданных в некотором базисе матрицами:

а)

$$\begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix}$$

б)

$$\begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}$$

8.4 Критерии диагонализируемости

Корни характеристического многочлена χ_{L_A} (говорят также: характеристические корни) составляют множество, несущее важную информацию о линейном операторе L_A. По понятным причинам, однако, не все характеристические корни равноправны.

Определение 32. *Множество всех собственных значений линейного оператора Л называют спектром этого оператора и обозначают символом* $\mathrm{Spec}L_A$ *(собственные значения считаются с их геометрическими кратностями). Аналогично говорят о спектре* $\mathrm{Spec}A$ *матрицы A. Точка спектра называется простой, если ей отвечает алгебраическая кратность* !. *Если все точки спектра простые, то и спектр называется простым.*

В случае алгебраически замкнутого поля, например, $F = \mathbb{C}$, характеристические корни совпадают с точками спектра, но в общем случае спектр может быть пуст, как, например, для оператора поворота на вещественной плоскости.

Лемма 2. *Собственные векторы, принадлежащие к различным собственным значениям, линейно независимы. Сумма* $\sum\limits_{\lambda \in \mathrm{Spec} L_A} V^\lambda$ *прямая (вообще говоря,* $\sum\limits_{\lambda \in \mathrm{Spec} L_A} V^\lambda$ *не совпадает с* V*).*

Доказательство. Пусть $\lambda_1, \ldots, \lambda_m$ — какие-то различные собственные значения, $V^{\lambda_1}, \ldots, V^{\lambda_m}$ — соответствующие собственные подпространства. Выберем в каждом V^{λ_i} по одному собственному вектору e_i. Нужно доказать их линейную независимость. Для $m = 1$ утверждение верно. Рассуждая по индукции относительно m и предполагая существование нетривиальной линейной зависимости $a_1 e_1 + a_2 e_2 + \ldots + a_m e_m = 0$, где, скажем, $a_1 \neq 0$, мы применим к обеим частям этого равенства оператор L_A. Так как $L_a e_i = \lambda_i e_i$, то $a_1\lambda_1 e_1 + a_2\lambda_2 e_2 + \ldots + a_m\lambda_m e_m = 0$. Умножая первое соотношение на λ_m и вычитая из него второе, приходим к линейной зависимости первых $m-1$ векторов: $a_1(\lambda_m - \lambda_1)e_1 + a_2(\lambda_m - \lambda_2)e_2 + \ldots + a_{m-1}(\lambda_m - \lambda_{m-1})e_{m-1} = 0$. По предположению индукции $a_i(\lambda_m - \lambda_i) = 0$, $i = l, \ldots, m-1$. Но $a_1 \neq 0$, $\lambda_i \neq \lambda_m$, $i < m \Rightarrow a_i(\lambda_m - \lambda_i) \neq 0$. Полученное противоречие доказывает наше утверждение.

По определению любой отличный от нуля вектор $e_i \in V^{\lambda_i}$ является собственным. Поэтому по доказанному $V^{\lambda_i} \bigcap \sum\limits_{i \neq j} V^{\lambda_j} = 0$. Это и значит, что сумма $\sum\limits_i V^{\lambda_i}$ прямая. □

Определение 33. *Линейный оператор* L_A *на* n*-мерном пространстве* V *называется диагонализируемым, если существует базис* (e_i)*, относительно которого матрица оператора принимает диагональный вид*

$$A = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}.$$

Теорема 17. *Линейный оператор* L_A *с простым спектром диагонализируем.*

Доказательство. Формулировка теоремы предполагает, что многочлен $\chi_A(t)$ имеет в основном поле F $n = \dim V$ различных корней $\lambda_i, \ldots, \lambda_n$, которым отвечают собственные векторы e_i, $i = 1, \ldots, n$. По лемме 2 эти

векторы линейно независимы. Значит, $V = \langle e_1, \ldots, e_n \rangle$, и так как $L_A e_i = \lambda_i e_i$, то $L_A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. $\square$

Простота спектра оператора является всего лишь достаточным условием его диагонализируемости. Например, идемпотентный оператор диагонализируем, хотя его спектр при $n \geq 2$ не будет простым.

Теорема 18. *Пусть L_A — линейный оператор на конечномерном векторном пространстве V над полем F. Для диагонализируемости L_A необходимо и достаточно выполнения следующих двух условий:*

1) все корни характеристического многочлена χ_A лежат в F;

2) геометрическая кратность каждого собственного значения L_A совпадает с его алгебраической кратностью.

Доказательство. Пусть выполнены условия 1), 2). Если $\lambda_1, \ldots, \lambda_m$ — различные корни многочлена χ_A, а $f_i, \ldots, f_m$ — их кратности, то

$$\dim V^{\lambda_i} = f_i, f_i + f_2 + \ldots + f_m = n. \tag{8.7}$$

По лемме 2 любая совокупность не равных одновременно нулю векторов $v_i \in V^{\lambda_i}$, $i = 1, \ldots, m$, линейно независима, так что

$$V^{\lambda_i} \bigcap (V^{\lambda_1} + \ldots + \hat{V}^{\lambda_i} + \ldots + V^{\lambda_m}) = 0 \tag{8.8}$$

Значит, сумма $V^{\lambda_1} + \ldots + V^{\lambda_m}$ прямая, а с учётом равенств (8.7) получаем

$$V = V^{lambda_1} + \ldots + V^{\lambda_m}. \tag{8.9}$$

Взяв за базис в V объединение базисов в V^{λ_i}, мы придём к собственному базису, т.е. к базису, состоящему из n линейно независимых собственных векторов оператора L_A. Его существование эквивалентно диагонализируемости L_A.

Обратно: пусть оператор L_A диагонализируем. Доказать самостоятельно. $\square$

Теорема 19. *Всякий комплексный (соответственно вещественный) линейный оператор L_A имеет одномерное (соответственно одномерное или двумерное) инвариантное подпространство.*

Доказательство. Так как характеристический многочлен χ_A имеет в $\mathbb{C}$ хотя бы один корень, то известный метод нахождения собственных векторов заведомо даст одномерное инвариантное подпространство исходного пространства V. В случае вещественного поля $\mathbb{R}$ рассмотрим минимальный многочлен $\mu_A(t)$ оператора L_A. Его коэффициенты лежат в $\mathbb{R}$. Если

$\mu_A(t)$ имеет вещественный корень a, то $\mu_A(t) = (t-a)g(t)$, $g(t) \in \mathbb{R}[t]$. Так как $g(L_A) \neq 0$ в силу минимальности $\mu_A(t)$, то $g(L_A)u \neq 0$ для некоторого вектора $u \in V$. Но $(L_A - aId)u = (L_A - aId)g(L_A)u = \mu_A(L_A)u = 0$, откуда $L_A u = au$, т.е. u — собственный вектор.

Предположим теперь, что L_A не имеет собственных векторов. Тогда по доказанному у $\mu_A(t)$ нет вещественных корней. Но по теореме о многочленах с вещественными коэффициентами мы имеем право записать $\mu_A(t) = (t^2 - at - b)h(t)$, $a, b \in \mathbb{R}$, $h(t) \in \mathbb{R}[t]$. Снова $v = h(A)u \neq 0$ для некоторого $u \in V$ и $L_A^2 v - aL_A v - bv = \mu_A(L_A)v = 0$. Получается, что $L_A^2 v = aL_A v + bv$, а так как $L_A v \neq \lambda v$ (одномерного инвариантного подпространства нет), то $L = \langle v, L_A v\rangle$ — двумерное инвариантное подпространство. □

Задача 16. *1. Выяснить, какие из следующих матриц можно привести к диагональному виду путём перехода к новому базису над полем* $\mathbb{R}$ *или над полем* $\mathbb{C}$*:*

а)

$$\begin{pmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{pmatrix}$$

б)

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Рассмотрим пространство V^* линейных функций на V. Тогда можно ввести как операцию $(\cdot,\cdot) : V^* \times V \to F$, так и

Определение 34. *Линейный оператор* L_A^* *на* V^**, заданный соотношением*

$$(L_A^* f, x) := (f, L_A x), \tag{8.10}$$

называют оператором, сопряжённым к L_A*.*

Теорема 20. *Всякий комплексный линейный оператор* L_A *на* V *обладает инвариантной гиперплоскостью.*

Доказательство. Пусть $\dim V = n$. Как мы знаем, $\dim \mathrm{Ker} f = n - 1$ для любой линейной функции $f \neq 0$ на V. Возьмём теперь в качестве f собственный вектор линейного оператора L_A^* на V^*. Он существует по теореме 19, и если λ — отвечающее ему собственное значение, то, как следует из определяющего равенства (8.10), $x \in \mathrm{Ker} f \Rightarrow 0 = \lambda(f, x) = (\lambda f, x) = (L_A^* f, x) = (f, L_A x) \Rightarrow L_A x \in \mathrm{Ker} f$. Это и означает, что $\mathrm{Ker} f$ — искомая гиперплоскость. □

8.5 Теорема Гамильтона-Кэли

Теорема 21. *Матрицу линейного оператора L_A всегда можно привести (в смысле подобия) к треугольному виду.*

Доказательство. Проще всего в этом убедиться рассуждением по индукции. По теореме 20 из пространство V содержит инвариантную относительно L_A гиперплоскость $U : L_A U \subset U$. По предположению индукции в U можно выбрать такой базис $(e_i, \dots, e_{n-i})$, что $L_A e_i = \lambda_i e_i + v_i$, $v_i \in \langle e_i, \dots, e_{n-i} \rangle$. Имеем $V = \langle U, e_n \rangle$, где e_n — произвольный, не содержащийся в U вектор. Пусть $L_A e_n = \lambda_n e_n + u$, $u \in U$. Таким образом, в базисе $(e_i, \dots, e_{n-i}, e_n)$ дейст- вие оператора L_A выражается матрицей требуемого вида

$$A = \begin{pmatrix} \lambda_1 & \dots & \dots & \dots \\ 0 & \lambda_2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \tag{8.11}$$

□

Теорема 22. *(Гамильтона—Кэли). Линейный оператор L_A и соответствующая ему матрица A (в любом базисе) аннулируются своим характеристическим многочленом $\chi_A(t)$ т.е. $\chi_A(A) = 0$.*

Доказательство. Так как это утверждение не зависит от выбора базиса, то естественно воспользоваться теоремой 21, с самого начала считая матрицу A в базисе $(e_1, \dots, e_n)$ имеющей треугольный вид (8.11).

Рассмотрим цепочку L_A-инвариантных подпространств $V = V_0 \supset V_1 \supset \dots \supset V_{n-1} \supset V_n = 0$, где $V_k = \langle e_i, \dots, e_{n-k-i}, e_{n-k} \rangle$. Так как $(L_A - \lambda_{n-k} Id) e_{n-k} \in V_{k+i}$, то $(L_A - \lambda_{n-k} Id) V_k \subset V_{k+1}$, и, стало быть, $\chi_A(A) = \prod_{i=1}^{n} (L_A - \lambda_i Id) V = (L_A - \lambda_1 Id) \cdots (L_A - \lambda_n Id) V_0 \subset (L_A - \lambda_1 Id) \cdots (L_A - \lambda_{n-1} Id) V_1 \subset \dots \subset (L_A - \lambda_1 Id) V_{n-1} = 0$.

Но $\chi_A(A) V = 0 \Leftrightarrow \chi_A(A) = 0$. □

Следствие 3. *Минимальный многочлен μ_A линейного оператора является делителем характеристического многочлена $\chi_A(t)$, делящимся на все линейные множители $t - \lambda$, $\lambda \in \mathrm{Spec}(L_A)$.*

Доказательство — упражнение.

8.6 ЖНФ

Определение 35. *1) Назовём (верхней) клеткой Жордана размера $m \times m$ (или порядка m), соответствующей собственному значению λ, матрицу*

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & \dots & 0 \\ 0 & \lambda & 1 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

2) Жордановой матрицей называется матрица, состоящая из диагональных блоков $J_{m_i}(\lambda_i)$ и нулей вне этих блоков:

$$J = \begin{pmatrix} J_{m_1}(\lambda_1) & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & J_{m_s}(\lambda_s) \end{pmatrix} \tag{8.12}$$

3) Жордановым базисом для линейного оператора $L_A : V \to V$ называется такой базис пространства V, в котором матрица оператора L_A является жордановой, или, как говорят, имеет жорданову нормальную форму (ЖНФ) $J(L_A)$.

4) Приведением квадратной матрицы A к жордановой нормальной форме называется решение уравнения в матрицах вида $X^{-1}AX = J(A)$, где X — (неизвестная) невырожденная матрица, а $J(A)$ — (неизвестная) жорданова матрица.

8.7 Корневые подпространства

Определение 36. *Множество векторов $V(\lambda) = \{v \in V | \exists k \in \mathbb{N}, (L_A - \lambda Id)^k v = 0\}$ называется корневым подпространством, соответствующим собственному значению $\lambda \in \mathrm{Spec}A$.*

Теорема 23. *Пусть $L_A : V \to V$ — линейный оператор с характеристическим многочленом $\chi_A(t) = \prod_{i=1}^{p}(t - \lambda_i)^{n_i},$, $\lambda_i \neq \lambda_j$ при $i \neq j$. Тогда $V = V(\lambda_i) \oplus \ldots \oplus V(\lambda_p)$ — прямая сумма корневых подпространств $V(\lambda_i)$, каждое из которых инвариантно относительно L_A и имеет размерность $\dim V(\lambda_i) = n_i$. Оператор $L_a - \lambda_i Id$, нильпотентный на $V(\lambda_i)$, действует невырожденным образом на подпространстве $V_i = V(\lambda_i) \oplus \ldots \oplus V(\lambda_{i-1}) \oplus V(\lambda_{i+1}) \oplus \ldots V(\lambda_p)$. Наконец, λ_i — единственное собственное значение оператора $L_A|_{V(\lambda_i)}$.*

Доказательство. Ни один из простых множителей $t-\lambda_k$ не может быть делителем одновременно всех многочленов $\chi_i(t) = \prod\limits_{j\neq i}(t-\lambda_j)^{n_j}$, $i = 1, \ldots, p$ и поэтому $GCD(\chi_1(t, \ldots, \chi_p(t)) = 1)$. Следовательно, найдутся многочлены $f_1(t), \ldots, f_p(t) \in \mathbb{C}[t]$, для которых

$$\sum_{i=1}^{p} \chi_i(t) f_i(t) = 1 \tag{8.13}$$

Подпространства $W_i = \chi_i(L_A) f_i(L_A) V = \{\chi_i(L_A) f_i(L_A) v | v \in V\}$, $1 \leq i \leq p$, инвариантны относительно L_A. Кроме того, $(A - \lambda_i Id)^{n_i} W_i = \chi_i(L_A) f_i(L_A) V = 0$ (поскольку по теореме 22 $\chi_i(L_A) = 0$), так что

$$W_i \subset V(\lambda_i). \tag{8.14}$$

Соотношение (8.13), переписанное в виде $Id = \sum\limits_{i=1}^{p} \chi_i(A) f_i(A)$ даёт нам разложение $V = \sum\limits_{i=1}^{p} W_i$ и тем более (ввиду включения (8.14)) $V = \sum\limits_{i=1}^{p} V(\lambda_i)$.

Предположим, что $v \in V(\lambda_i) \bigcap V_i$, где, как и в формулировке теоремы, $V_i = \sum\limits_{i\neq j} V(\lambda_j)$. Тогда $(L_A - \lambda_i Id)^n v = 0$, а так как $v = \sum\limits_{i\neq j} v_j$ и $(L_A - \lambda_j Id)^n v_j = 0$, то и $(\prod\limits_{j\neq i}(L_A - \lambda_j Id)^n) v = 0$. Но из взаимной простоты многочленов $(t-\lambda_i)^n$, $c(t) = \prod\limits_{j\neq i}(t-\lambda_j)^n$ следует существование $a(t)$, $b(t)$, для которых $a(t)(t-\lambda_i)^n + b(t)c(t) = 1$. Получаем $v = a(L_A)(L_A - \lambda_i Id)^n v + b(L_A)(\prod\limits_{j\neq i}(L_A - \lambda_j Id)^n) v = 0$, т.е. пространства $V(\lambda_i)$ и V_i не пересекаются. Значит, мы имеем разложение

$$V = V(\lambda_1) \oplus \ldots \oplus V(\lambda_p) \tag{8.15}$$

в прямую сумму L_A-инвариантных подпространств.

Из включения (8.14) и из разложения (8.15) непосредственно вытекает, что $W_i = V(\lambda_i)$. Таким образом, для $V(\lambda_i)$ получено выражение $V(\lambda_i) = \chi_i(L_A) f_i(L_A) V$, где $\chi_i(L_A)$, $f_i(L_A)$ — многочлены из тождества (8.13). В частности, $(L_A - \lambda_i Id)^n V(\lambda_i) = 0$.

Минимальным многочленом для L_A на $V(\lambda_i)$ будет некоторый делитель многочлена $(t-\lambda_i)^{n_i}$. Отсюда следует, во-первых, что λ_i — единственное собственное значение оператора $L_A|_{V(\lambda_i)}$. Далее, в ба- базисе, являющемся объединением базисов пространств $V(\lambda_i)$, оператор L_A имеет матрицу

$$A = \begin{pmatrix} A_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & A_p \end{pmatrix},$$

где A_i — матрица порядка $m_i = \dim V(\lambda_i)$ с единственным собственным значением λ_i и характеристическим многочленом $\chi_{A_i}(t) = (t-\lambda_i)^{m_i}$, $m_i \le n_i$. Так как $\chi_a(t) = \prod_{i=1}^{p} \chi_{A_i}(t)$, то $n = n_i + \cdots + n_p$ и $m_i = n_i$.

Осталось доказать невырожденность ограничения $(L_A - \lambda_i Id)|_{V_i}$. Но в противном случае $(\mathrm{Ker}(L_A - \lambda_i Id) \bigcap V_i \neq 0$ и $L_a v - \lambda_i v = 0$ для некоторого $0 \neq v \in V_i$. Однако на V_i характеристическим многочленом для L_a является $\chi_i(t) = \prod_{j \neq i}(t-\lambda_j)^{n_j}$, и λ_i собственным значением быть не может. □

8.8 ЖНФ нильпотентного оператора

Определение 37. *Линейная оболочка $F(L_A)v = \langle v, L_A v, L_A^2 v, \ldots, L_A^{n'-1} v\rangle$ называется циклическим подпространством, ассоциированным с оператором L_A индекса нильпотентности n и вектором v. Предполагается, что $n' \le n$ — наименьшее натуральное число, для которого $L_A^{n'} v = 0$.*

Теорема 24. *Жорданова нормальная форма $J(A)$ нильпотентной матрицы A существует (основное поле F произвольно*

Доказательство. Из определения 37 видно, что всякому циклическому подпространству отвечает клетка Жордана. Нам нужно показать, что векторное пространство V, на котором действует нильпотентный оператор L_A с матрицей A, разлагается в прямую сумму надлежащим образом выбранных циклических подпространств.

По теореме 21 матрица A приводится к верхнему треугольному виду с нулями по диагонали. Это значит, что линейная оболочка U первых $n-1$ базисных векторов инвариантна относительно L_A. По определению $L_A V \subset U$, а по предположению индукции в U можно выбрать жорданов базис для A, или, что то же самое,

$$U = F[L_A]e_1 \oplus \ldots \oplus F[L_A]e_s, \tag{8.16}$$

$$F[L_A]e_i = \langle e_i, L_A e_i, L_A^2 e_i, \ldots, L_A^{m_i-1} e_i\rangle, L_A^{m_i} e_i = 0$$

Без ограничения общности считаем

$$m_1 \ge m_2 \ge \ldots \ge m_s \tag{8.17}$$

Далее, $V = \langle v, U\rangle$, $L_A v \in U$ для любого вектора v, не содержащегося в U, так что $L_A v = \sum_i \alpha_i e_i + L_A u$, $u \in U$. Заменяя v на $v' = v - u$, будем иметь $V = \langle v', U\rangle$, $L_A v' = \sum_{i=1}^{s} \alpha_i e_i$.

Если $\alpha_i = 0$, $1 \le i \le s$, то к клеткам Жордана $J_{m_i}(0), \ldots, J_{m_s}(0)$ добавится $J_1(0)$, отвечающая циклическому подпространству $\langle v' \rangle$, то есть $A \sim J(A) = \mathrm{diag}(J_{m_i}(0), \ldots, J_{m_s}(0), J_1(0))$. Остаётся рассмотреть случай, когда $\alpha_1 = \ldots \alpha_{r-1} = 0$, $L_A v' = \sum\limits_{i=r}^{s} \alpha_i e_i$, $\alpha_r \neq 0$ для некоторого индекса $r \ge 1$. Удобно положить $e_i^{'} = e_i$, $i \neq r$, $e_r^{'} = \frac{1}{\alpha_r} v'$, $\beta_i = \frac{\alpha_i}{\alpha_r}$.

Тогда $L_A e_r^{'} = e_r + \sum\limits_{i=r+1}^{s} \beta_i e_i := f_r$. В соответствии с упорядочением (8.17) $L_A^{m_r} f_r = 0$, а так как сумма (8.16) прямая, то $L_A^{m_r - 1} f_r \neq 0$, какие бы ни были коэффициенты β_i. Кроме того, простое рассуждение показывает, что сумма $\sum i \neq r f[L_A] e_i^{'} + F[L_A] f_r$ также является прямой и совпадает с U.

Но теперь циклическое подпространство $F[L_A] f_r$ расширяется за счёт вектора $e_r^{'} \notin U$: $F[L_A] f_r \subset f[L_A] e_r^{'}$, и мы имеем прямую сумму $V = \bigoplus\limits_{i=1}^{s} f[L_A] e_i^{'}$, отвечающую набору индексов $m_1^{'}, \ldots, m_s^{'}$, где $m_i^{'} = m_i$, $i \neq r$, $m_r^{'} = m_r + 1$. В свою очередь $A \sim \mathrm{diag}(J_{m_1^{'}}(0), \ldots, J_{m_s^{'}}(0))$. (число клеток Жордана сохранилось прежним, но размер одной клетки увеличился на 1). Последовательность $(m_1^{'}, \ldots, m_s^{'})$, вообще говоря, не упорядочена, но этого всегда можно добиться путём переобозначения векторов $e_i^{'}$. Таким образом, существование жорданова базиса для нильпотентного оператора L_A доказано. $\square$

8.9 Единственность

Приступая к доказательству единственности, укажем заодно практическое правило для приведения произвольной матрицы A порядка n к жордановой нормальной форме. Для этого нужно уметь находить число $N(m, \lambda)$ жордановых клеток $J_m(A)$ порядка m, отвечающих собственному значению λ матрицы A. Сопоставим обычным образом матрице A оператор L_A, действующий на n-мерном векторном пространстве V, и разложим V в прямую сумму

$$V = V(\lambda) \oplus V' \tag{8.18}$$

где $V(\lambda) = \bigoplus\limits_{j=1}^{s} \langle e_j, (L_A - \lambda Id) e_j, \ldots, (L_A - \lambda Id)^{m_j - 1} e_j \rangle$, $V' = \sum\limits_{\lambda' \neq \lambda} V(\lambda')$. Будем подсчитывать ранг $r_t = \mathrm{rank}(A - \lambda Id)^t$ матрицы $(A - \lambda Id)^t$, или, что то же самое, размерность пространства $(L_A - \lambda Id)^t V$. Эта размерность, конечно, не зависит от выбора базиса в V. Каждое из пространств в раз-

ложении (8.18) инвариантно относительно $(L_A-\lambda Id)^t$, поэтому $\dim(L_A-\lambda Id)^tV=\sum\limits_j \dim(L_A-\lambda Id)^t\mathbb{C}[L_A]e_j+\dim(L_A-\lambda Id)^tV'$.

Пусть для определённости $m_1\le m_2\le\ldots\le m_s$. Если $m_j\le t$, то $(L_A-\lambda Id)^t\mathbb{C}[L_A]e_j=0$. При $m_j>t$ имеем $(L_A-\lambda Id)^t\mathbb{C}[L_A]e_j=\langle(L_A-\lambda Id)^te_j,(L_A-\lambda Id)^{t+1}e_j,\ldots,(L_A-\lambda Id)^{m_j-1}e_j\rangle$ так что $\dim(L_A-\lambda Id)^t\mathbb{C}[L_A]e_j=m_j-t$. На V оператор $(L_A-\lambda Id)$ невырожден (теорема 21), поэтому $\dim(L_A-\lambda Id)^tV'=\dim V'$. Получаем $r_t=\sum\limits_{m_j>t}(m_j-t)+\dim V'$, откуда $r_t-r_{t+i}=\sum\limits_{m_j>t}(m_j-t)-\sum\limits_{m_j>t+1}(m_j-t-1)=\sum\limits_{m_j>t}(m_j-t)-\sum\limits_{m_j>t+1}(m_j-t)+\sum\limits_{m_j>t+1}1=\sum\limits_{m_j=t+1}1+\sum\limits_{m_j>t+1}1=N(t+1,\lambda)+N(t+2,\lambda)+\ldots$. Следовательно, $r_{m-i}-r_m-(r_m-r_{m+i})=(N(m,\lambda)+N(m+1,\lambda)+\ldots)-(N(m+1,\lambda)+N(m+2,\lambda)+\ldots)=N(m,\lambda)$, и мы получаем окончательную формулу

$$N(m,\lambda)=r_{m-i}-2r_m+r_{m+i},\qquad(8.19)$$

$$m\ge 1,\ r_t=\operatorname{rank}(L_A-\lambda Id)^t,\ r_0=n.$$

Заметим, что r_t — инвариант матрицы A (т.е. число, определяемое классом подобия матрицы A). Значит, формулой (8.19) устанавливается также единственность жордановой формы $J(A)$.

Задача 17. *Найти ЖНФ матриц, найти базис пространств, соответствующих собственным значениям*

a)
$$\begin{pmatrix}1&1&-1\\-3&-3&3\\2&-2&1\end{pmatrix}.$$

b)
$$\begin{pmatrix}1&-1&2\\3&-3&6\\2&-2&3\end{pmatrix}.$$

c)
$$\begin{pmatrix}9&22&-6\\-1&-4&1\\8&16&-5\end{pmatrix}.$$

d)
$$\begin{pmatrix}3&1&0&0\\-4&-1&0&0\\7&1&2&1\\-17&-6&-1&0\end{pmatrix}.$$

e)

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Глава 9

Квадратичные формы

9.1 Определение

Рассмотрение симметричных билинейных форм приводит к следующему важному понятию, которое естественным образом возникает в разных разделах математики.

Определение 38. *Квадратичной формой на конечномерном векторном пространстве V над $\mathbb{R}$ (или любым другим полем) называется функция $q : V \to \mathbb{R}$, обладающая двумя свойствами:*
i) $q(-v) = q(v) \forall v \in V$;
ii) отображение $F : V \times V \to \mathbb{R}$, определённое формулой

$$F(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)), F) \tag{9.1}$$

является билинейной формой на V (очевидно, симметричной). Её ранг называется также рангом q: $\mathrm{rank}q = \mathrm{rank}F$. *Говорят ещё, что симметричная билинейная форма F, определённая формулой (9.1), получается из q поляризацией или что F — билинейная форма, полярная к квадратичной форме q.*

Пусть теперь f — произвольная симметричная билинейная форма на V. Положив

$$q_f(x) = f(x, x) \tag{9.2}$$

мы получим функцию $q_f : V \to \mathbb{R}$, удовлетворяющую условиям i), ii) в определении квадратичной формы, поскольку $f(-x, -x) = f(x, x)$ и $f(x, y) = 1/2(f(x + y, x + y) - f(x, x) - f(y, y))$.

Теорема 25. *Каждая квадратичная форма q однозначно восстанавливается по своей полярной форме f; другими словами, $q = q_f$.*

Доказательство. Положим в (9.1) $y = -x$. Тогда $q(x) = f(x,x)+1/2q(0)$. Так как f — билинейная форма, то $f(0,0) = 0$. Поэтому при $x = 0$ имеем $q(0) = 1/2q(0)$, т.е. $q(0) = 0$. Значит, $q(x) = f(x,x)$. □

Определение 39. *Матрицей квадратичной формы $q = q_f$ относительно базиса $(e_i, \ldots, e_n)$ пространства V называется матрица F билинейной формы f, полярной к q, то есть $F = (f_{ij})$, где $f_{ij} = \frac{1}{2}(q(e_i + e_j) - q(e_i) - q(e_j))$, $i, j = 1, 2, \ldots, n$.*

Любой симметричной матрице $F = (f_{ij})$ в свою очередь отвечает квадратичная форма q, заданная соотношением

$$q(x) = x^t F x = \sum_{i,j=1}^{n} f_{ij} x_i x_j, x \in V \tag{9.3}$$

Таким образом, в соответствии с названием квадратичная форма есть однородная квадратичная функция координат $x_1, \ldots, x_n$ вектора $x = \sum\limits_{i=1}^{n} x_i e_i$.

Определение 40. *Говорят, что квадратичная форма q имеет в базисе $(e_1, \ldots, e_n)$ пространства V канонический или диагональный вид, если для каждого вектора $x \in V$ значение $q(x)$ вычисляется по формуле*

$$q(x) = \sum_{i=1}^{n} f_{ii} x_i^2 \tag{9.4}$$

Базис $(e_1, \ldots, e_n)$ при этом называется каноническим базисом для q.

9.2 Существование канонического вида квадратичной формы

Вопрос о возможности выбора базиса, в котором данная форма принимала бы канонический вид, имеет важное теоретическое и прикладное значение.

Теорема 26. *Для всякой симметричной билинейной формы f на V существует канонический базис.*

Доказательство. По индукции по размерности V. При $n = 1$ утверждение очевидно.

Если $f(x,y)=0$ для всех $x,y\in V$ (т.е. $f=0$), то теорема очевидна: любой базис годится. Пусть $f\neq 0$, тогда отлична от нуля и соответствующая квадратичная форма (теорема 25). Пусть e_1 — такой вектор, что $f(e_i,e_i)=q(e_i)\neq 0$. Тогда линейная функция $f_1: x\mapsto f(x,e_i)$ отлична от нуля. Тогда линейное подпространство $L=\mathrm{Ker} f_1=\{x\in V|f_1(x)=0\}$ имеет размерность $n-1$, т.е. является гиперплоскостью. По предположению индукции L обладает базисом $(e_2,\ldots,e_n)$, в котором матрица формы f, ограниченной на L, диагональна, т.е. $f(e_i,e_j)=0$ при $i\neq j$, $i,j=2,\ldots,n$. Так как по построению $f(e_i,e_1)=0$, $i=2,3,\ldots,n$, то мы получаем свойства $f(e_i,e_j)=0$ при $i\neq j$, характеризующие канонический базис $(e_1,\ldots,e_n)$, если только система векторов $(e_1,\ldots,e_n)$ линейно независима. Предположив противное, мы в любом соотношении $\alpha_1 e_i+\ldots+\alpha_n e_n=0$ имели бы коэффициент $\alpha_1\neq 0$, поскольку $(e_1,\ldots,e_n)$ — базис в L. Но в таком случае $e_i=\sum\limits_{i=2}^{n}\beta_i e_i$ и $0\neq f_1(e_1)=f_1(\sum\limits_{i=2}^{n}\beta_i e_i)=\sum\limits_{i=2}^{n}\beta_i f_1(e_i)=0$ — противоречие, доказывающее теорему. □

Следствие 4. *Пусть на векторном пространстве V размерности n над полем $\mathbb{F}$ задана квадратичная форма q ранга $r\leq n$. Тогда в V существует базис $(e_1,\ldots,e_n)$, в котором q принимает канонический вид.*

Следствие 5. *Для любой симметричной матрицы F существует такая невырожденная матрица A, что A^tFA — диагональная матрица того же ранга, что и F. Другими словами, всякая симметричная матрица конгруэнтна диагональной.*

9.3 Метод Лагранжа приведения квадратичной формы к каноническому виду

Рассмотрим квадратичную форму $q(x)=\sum\limits_{i,j=1}^{n} f_{ij}x_i x_j$. Выделим все члены, содержащие координату x_1.

$$q(x_1,\ldots,x_n)=f_1 1x_1^2+2f_{12}x_1x_2+\ldots+2f_{1n}x_1x_n+\sum_{i,j=2}^{n}f_{ij}x_ix_j.$$

Пусть $f_{11}\neq 0$. Тогда

$$q(x_1,\ldots,x_n)=\frac{1}{f_{11}}(\sum_{i=1}^{n}f_{1i}x_i)^2+\sum_{i,j=2}^{n}f'_{ij}x_ix_j.$$

Положим $x_1^{'} = \sum_{i=1}^{n} f_{1i}x_i$. Тогда

$$q(x_1^{'}, \ldots, x_n) = \frac{1}{f_{11}} x_1^{'2} + q'(x_2, \ldots, x_n).$$

И т.д. Если $f_{11} = 0$ рассмотрим замену переменных $x_1^{'} = x_1 - x_2$, $x_2^{'} = x_1 + x_2$. В такой системе координат $f_{11}^{'} \neq 0$.

9.4 Нормальный вид квадратичной формы

Определение 41. *Говорят, что квадратичная форма q, значения которой вычисляются по формуле*

$$q(x) = \sum_{i=1}^{s} x_i^2 - \sum_{i=s+1}^{r} x_i^2$$

имеет нормальный вид.

Следствие 6. *Всякая квадратичная форма q на вещественном векторном пространстве V приводится к нормальному виду.*

Кроме ранга r у квадратичной формы q на векторном пространстве V над $\mathbb{F}$ появилась еще одна числовая характеристика — количество s коэффициентов 1 в её нормальном виде.

Теорема 27. *(закон инерции). Пусть q — квадратичная форма на n-мерном векторном пространстве V над $\mathbb{F}$. Тогда целые числа r и s, $s \leq r \leq n$, входящие в нормальный вид, зависят только от q.*

Доказательство. Инвариантность r нам известна, так что нужно лишь убедиться в инвариантности (независимости от выбора канонического базиса) числа s. Предположим, что в каком-то другом базисе $(e_1^{'}, \ldots, e_n^{'})$ форма q имеет нормальный вид $q(x) = x_1^{'2} + \ldots + x_t^{'2} - x_{t+1}^{'2} - \ldots - x_r^{'2}$. с t положительными членами. При $t \neq s$ без ограничения общности считаем $t < s$. Рассмотрим в V подпространства $L =< e_i, \ldots, e_s >$ $L' =< e_{t+1}^{'}, \ldots, e_n^{'}$. Так как $\dim(L + L') \leq \dim V = n$, то $\dim(L \bigcap L') = \dim L + \dim L' - \dim(L + L') \geq s + (n - t) - n = s - t > 0$. Следовательно, существует ненулевой вектор $x \in (L \bigcap L') : 0 \neq x = x_1 e_i + \ldots + x_s e_s = x_{t+1}^{'} e_{t+1}^{'} + \ldots + x_n^{'} e_n$. С одной стороны $q(x) = x_1^2 + \ldots + x_s^2 > 0$. В то же время $q(x) = -x_{t+1}^{'2} - \ldots - x_n^{'2} < 0$. Полученное противоречие устраняется только в случае $s = t$. $\square$

Ввиду теоремы 27 для числовых инвариантов формы используются специальные термины.

Определение 42. *Ранг вещественной квадратичной формы называется также её индексом инерции, число s — положительным индексом инерции, число $r-s$ — отрицательным индексом инерции. Под сигнатурой формы понимают либо пару $(s, r-s)$, либо разность $2s-r$ между числом положительных и числом отрицательных квадратов.*

Определение 43. *Невырожденная квадратичная форма $q : V \times V \to \mathbb{F}$ называется положительно (соответственно отрицательно) определённой или просто положительной (отрицательной), когда $q(x) > 0$ ($q(x) < 0$) для любого вектора $x \neq 0$. Форма q называется положительно полуопределённой (или неотрицательной), если $q(x) \geq 0$ для всех $x \in V$. Наконец, форма q неопределённая, если она принимает как положительные, так и отрицательные значения.*

Следствие 7. *Любая положительно определённая матрица F имеет вид $F = A^{tr}A$, где A — вещественная невырожденная матрица. Верно и обратное: всякая вещественная матрица вида $A^{tr}A$ положительно определена.*

Рассмотрим следующие полезные объекты, носящие название главных миноров матрицы F:

$$\Delta_1 = f_{11}, \Delta_2 = \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix}, \ldots, \Delta_i = \begin{vmatrix} f_{11} & \ldots & f_{1i} \\ \ldots & \ldots & \ldots \\ f_{i1} & \ldots & f_{ii} \end{vmatrix}$$

Пусть $\Delta_0 = 1$.

Теорема 28. *(метод Якоби). Пусть q — квадратичная форма на V с матрицей F, все главные миноры которой отличны от нуля. Тогда существует базис $(e_i', \ldots, e_n')$ пространства V, в котором $q(x)$ принимает канонический вид*

$$q(x) = \frac{\Delta_0}{\Delta_1} x_1'^2 + \ldots + \frac{\Delta_{n-1}}{\Delta_n} x_n'^2.$$

Доказательство. Доказательство по индукции по n. Пусть $(e_i, \ldots, e_n)$ — первоначальный базис пространства V. Рассмотрим $(n-1)$-мерное подпространство $L =< e_i, \ldots, e_{n-1} >$. Пусть $q|_L$ — ограничение q на L. Матрица F формы q получается из F вычёркиванием последней строки и последнего столбца, поэтому её главными минорами будут $\Delta_1' = \Delta_1$,

$\ldots$, $\Delta_{n-1}^{'} = \Delta_{n-1}$. Все они по условию отличны от нуля. Выберем в L базис, в котором $q(x)$, $x \in L$, принимает вид

$$q|_L(x) = \frac{\Delta_0}{\Delta_1} x_1^{'2} + \ldots + \frac{\Delta_{n-2}}{\Delta_{n-1}} x_{n-1}^{'2}.$$

Тогда $f(e_i^{'}, e_j^{'}) = 0$, $i \neq j$ и $f(e_i^{'}, e_i^{'}) = \frac{\Delta_{i-1}}{\Delta_i}$. Рассмотрим ситсему из $n-1$ линейного уравнения:

$$f(x, e_i^{'}) = 0, \ldots, f(x, e_{n-1}^{'}) = 0.$$

Эта система имеет решение, обозначим его через $e_n^{'}$ и нормируем так, чтобы матрица перехода A от $(e_i, \ldots, e_n)$ к $(e_i^{'}, \ldots, e_n^{'})$ имела определитель $\det A = \frac{1}{\det F}$.

Пусть F' — матрица формы q в базисе $(e_i^{'}, \ldots, e_n^{'})$. Тогда $f(e_i^{'}, e_j^{'}) = 0$, $i \neq j$ и

$$f(e_n^{'}, e_n^{'}) = \Delta_{n-1} \frac{\Delta_0}{\Delta_1} \cdots \frac{\Delta_{n-2}}{\Delta_{n-1}} f(e_n^{'}, e_n^{'}) = \Delta_{n-1} \prod_{i=1}^{n} f(e_i^{'}, e_i^{'}) = \Delta_{n-1} \det F' = \Delta_{n-1} \det(A^{tr} F A) =$$

□

Задача 18. *1. Найти нормальный вид квадратичных функций:*

$$x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$$

$$x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 4x_1x_3 + 2x_2x_3$$

$$x_1^2 - 3x_3^2 - 2x_1x_2 + 2x_1x_3 - 6x_2x_3$$

Глава 10

Комплексные числа

10.1 Определение множества комплексных чисел

Очевидно, что далеко не каждое линейное уравнение с натуральными коэффициентами имеет целое неотрицательное решение. С введением целых отрицательных чисел область решения таких уравнений была явно расширена. Решение любого уравнения первого порядка с целыми коэффициентами стало возможно лишь с появлением рациональных чисел. Исследование простейшего квадратного уравнения $x^2 - 2 = 0$ привело в конечном итоге к появлению полной системы действительных чисел. В настоящее время не все квадратные уравнения с действительными коэффициентами решаются в школьном курсе математики. Если дискриминант уравнения отрицательный, то по школьным учебникам такое уравнение решений не имеет. Самым простым среди квадратных уравнений, не имеющих действительных корней, является уравнение

$$x^2 + 1 = 0 \tag{10.1}$$

Поставим перед собой следующую задачу: построить новую систему чисел, которая, во-первых, содержала бы корень уравнения (10.1) и, во-вторых, являлась бы алгебраическим расширением системы действительных чисел. Второе условие в этой задаче означает, что новая система чисел должна содержать все действительные числа как подмножество и все числовые операции для новой системы, если они применяются к действительным числам, должны совпадать с известными операциями над действительными числами. Обозначим через $\mathbb{R}$ множество всех действительных чисел. Буквой $\mathbb{K}$ обозначим множество всех точек плоскости и будем рассматривать эти точки как элементы новой числовой системы.

Выберем на плоскости декартову систему координат. Будем считать, что по оси абсцисс располагаются действительные числа и при этом начало координат совпадает с числом 0. Каждая точка плоскости теперь однозначно определяется своими координатами, т. е. парой действительных чисел.

Таким образом, $\mathbb{K} = \{(\alpha, \beta) | \alpha, \beta \in \mathbb{R}\}$. Новая числовая система будет полностью построена, когда будут определены все основные операции для ее элементов, и, строго говоря, только после этого элементы $\mathbb{K}$ можно называть числами. Понимая это, до окончательного построения числовой системы формально назовем точки плоскости, а значит, все элементы $\mathbb{K}$, комплексными числами. Очевидно, что два комплексных числа равны тогда и только тогда, когда равны их соответствующие координаты. Очевидно, что каждое действительное число α как точка плоскости имеет координаты $(\alpha, 0)$. Таким образом, $\mathbb{R} \subset \mathbb{K}$.

Определение 44. *Пусть даны два комплексных числа $a = (\alpha, \beta)$, $b = (\gamma, \delta)$. Под суммой $a + b$ будем понимать такое комплексное число c, координаты которого находятся по следующему правилу: $c = (\alpha+\gamma, \beta+\delta)$*

Задача 19. $(\mathbb{K}, +)$ — *коммутативная группа по сложению.*

Определение 45. *Пусть $a = (\alpha, \beta)$, $b = (\gamma, \delta)$. Под произведением $a \cdot b = ab$ двух комплексных чисел a и b будем понимать такую точку c, координаты которой находятся по следующему правилу: $c = ab = (\alpha\gamma - \beta\delta, \alpha\delta + \beta\gamma)$.*

Задача 20. $(\mathbb{K} \setminus \{(0, 0)\}, \cdot)$ — *коммутативная группа по умножению.*

Задача 21. $(\mathbb{K}, +, \cdot)$ — *поле.*

Теперь осталось выяснить, содержит ли $\mathbb{K}$ корень уравнения (10.1). Обозначим через i комплексное число $(0, 1)$. Тогда $i^2 = (0, 1)(0, 1) = (-1, 0)$. Выше было условлено не отличать действительное число -1 от комплексного числа $(-1, 0)$, поэтому $i^2 = -1$. Очевидно, что число i является корнем уравнения (10.1) и поставленная в начале этого параграфа задача решена полностью.

Далее перейдем к другой, более удобной форме записи комплексных чисел. Пусть (α, β) — произвольное комплексное число. Очевидно, $(\alpha, \beta) = (\alpha, 0) + (0, \beta)$. Но $(0, \beta) = (\beta, 0)(0, 1)$. Учитывая, что $(\alpha, 0) = \alpha$, $(\beta, 0) = \beta$ и $(0, 1) = i$, окончательно получаем

$$(\alpha, \beta) = \alpha + \beta i.$$

Если комплексное число записано в виде $a = \alpha + \beta i$ либо $a = \alpha + i\beta$, то любую из этих форм записи комплексного числа будем называть алгебраической. По сложившейся терминологии число i называется мнимой единицей, α называется действительной, а $i\beta$ — мнимой частями комплексного числа. Плоскость, точки которой использованы для построения множества комплексных чисел, называют комплексной плоскостью) оси абсцисс и ординат в выбранной системе координат называют соответственно действительной и мнимой осями.

10.2 Тригонометрическая форма записи комплексных чисел

Пусть $a = (\alpha, \beta)$ — произвольное комплексное число. Соединим начало координат с точкой $A(\alpha, \beta)$ и длину полученного отрезка обозначим r. Далее угол между положительным направлением оси абсцисс и направлением из начала координат на эту точку обозначим ϕ. Из прямоугольного треугольника OAB имеем $\alpha^2 + \beta^2 = r^2$, $\alpha = r\cos\phi$, $\beta = r\sin\phi$. При этом, очевидно, $r = +\sqrt{\alpha^2 + \beta^2}$. Подставляя α и β в формулу $a = \alpha + \beta i$, получаем тригонометрическую форму записи комплексного числа a : $a = r(\cos\phi + r\sin\phi)$.

Определение 46. *Число r называется модулем, а угол ϕ — аргументом комплексного числа a.*

Для модуля и аргумента имеют место следующие обозначения: $r = |a|$, $\phi = \arg(a)$. Аргумент числа a считается положительным, если угол отсчитывается против часовой стрелки, и отрицательным — в противном случае. При этом любой из углов $\phi + 2k\pi$, где $k \in \mathbb{Z}$, также считается аргументом числа a. Аргумент не определен лишь для числа $0 = (0, 0)$, но это число вполне определяется равенством $|0| = 0$.

Задача 22. *Модуль произведения двух комплекс- комплексных чисел равен произведению модулей сомножителей, а аргумент произведения равен сумме аргументов сомножителей.*

Задача 23. *Модуль частного двух комплексных чисел равен частному от деления модуля делимого на модуль делителя, а аргумент частного равен разности аргумента делимого и аргумента делителя.*

10.3 Сопряженные числа

Определение 47. *Пусть дано комплексное число $a = \alpha + i\beta$. Тогда число $a = \alpha - i\beta$ будем называть сопряженным с a и обозначать $\overline{a}$. Замена знака на противоположный перед мнимой частью комплексного числа называется операцией сопряжения. Таким образом,*

$$\overline{a} = \overline{\alpha + i\beta} = \alpha - i\beta$$

По определению, $\overline{\overline{a}} = a$ и поэтому числа a и $\overline{a}$ сопряжены друг с другом. Очевидно, что если a — действительное число, то $\overline{a} = a$. Следовательно, всегда имеется не более двух сопряженных друг другу чисел.

Теорема 29. *Если число a некоторым образом выражено через комплексные числа $b_1, b_2, \ldots, b_n$ при помощи сложения, умножения, вычитания и деления, то, заменяя в этом выражении все числа b_i их сопряженными, мы получим число, сопряженное с a.*

Доказательство — упражнение.

Пример 10. *Найти модуль и аргумент числа $z = 1 + i$.*

$|z| = \sqrt{1^2 + 1^2} = \sqrt{2}$. *Если $\phi = \arg(z)$, то $\tan\phi = 1$. Поскольку число z находится в первой четверти, то $\arg z = \pi/4$.*

Пример 11. *Записать число $z = l + i\sqrt{3}$ в тригонометрической форме.*

$|z| = \sqrt{1+3} = 2$; *если $\phi = \arg(z)$, то $\tan\phi = \sqrt{3}$. Тогда $\arg(z) = \pi/3$. Следовательно, $z = 2(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3})$.*

Пример 12. *Используя тригонометрическую форму записи, выполнить действия: $(1+i)(l+i\sqrt{3})$, $\frac{(1+i)}{(l+i\sqrt{3})}$. Результаты записать в тригонометрической форме.*

$$(1+i)(l+i\sqrt{3}) = (\sqrt{2}(\cos\frac{\pi}{4}+i\sin\frac{\pi}{4}))(2(\cos\frac{\pi}{3}+i\sin\frac{\pi}{3})) = 2\sqrt{2}(\cos(\frac{\pi}{4}+\frac{\pi}{3})+i\sin(\frac{\pi}{4}+\frac{\pi}{3})) = 2\sqrt{}$$

$$\frac{(1+i)}{(l+i\sqrt{3})} = \frac{\sqrt{2}(\cos\frac{\pi}{4}+i\sin\frac{\pi}{4})}{2(\cos\frac{\pi}{3}+i\sin\frac{\pi}{3})} = \frac{\sqrt{2}}{2}(\cos(\frac{\pi}{4}-\frac{\pi}{3})+i\sin(\frac{\pi}{4}-\frac{\pi}{3})) = \frac{\sqrt{2}}{2}(\cos(-\frac{\pi}{12})+i\sin(-\frac{\pi}{12}$$

Глава 11

Многочлены

В школьном курсе математики рассматриваются выражения вида $a_k x^k$, где x — переменная (неизвестное), которая может принимать любые действительные значения, а a_k — числовой коэффициент. При этом степень k — любое, в том числе и отрицательное, целое число. Такие выражения в элементарной алгебре называют одночленами. В настоящей главе рассматриваются выражения, которые являются формальными конечными суммами одночленов от одного неизвестного x, причем все степени x — целые неотрицательные. Если некоторое такое выражение традиционно обозначить через $f(x)$, то его, после приведения подобных относительно одинаковых степеней x, всегда можно записать в следующем виде:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0. \tag{11.1}$$

Определение 48. *Любое выражение вида (11.1), где x — неизвестное, а $a_n, a_{n-1}, \ldots, a_0$ — числовые коэффициенты, будем называть многочленом от x. Под степенью многочлена будем понимать наивысшую степень его неизвестного x, входящую в многочлен с ненулевым коэффициентом.*

Так, если в правой части (11.1) коэффициент $a_n \neq 0$, то $f(x)$ — многочлен степени n, при этом a_n — старший коэффициент, $a_n x^n$ — старший член многочлена $f(x)$, a_0 — свободный член. Всюду далее при записи многочленов в виде (11.1) будем считать, что старший коэффициент отличен от нуля. Если же потребуется к записи многочлена приписать несколько слагаемых с нулевыми коэффициентами, то это будет отмечено специально. Для сокращенной записи многочленов будут употребляться символы $f(x)$, $g(x)$, $\phi(x)$, $\psi(x)$ и т. п. В этой главе рассматриваются многочлены с комплексными коэффициентами от неизвестного x, которая может принимать любые комплексные значения.

Обозначим множество многочленов от одной переменной x с коэффициентами из поля F через $F[x]$.

Определение 49. *Два многочлена $f(x)$ и $g(x)$ будут считаться равными (или тождественно равными), если равны их коэффициенты при одинаковых степенях неизвестного.*

Определим основные операции над многочленами. Пусть $f(x)$ и $g(x)$ — два многочлена с комплексными коэффициентами. Для удобства запишем эти многочлены по возрастающим степеням неизвестного: $f(x) = a_0+a_1x+\ldots+a_{n-1}x^{n-1}+a_nx^n$, $a_n \neq 0$; $g(x) = b_0+b_1x+\ldots+b_{m-1}x^{m-1}+b_mx^m$, $b_m \neq 0$. Пусть $m \leq n$. Тогда $g(x)$ можно представить в виде $g(x) = b_0 + b_1x + \ldots + b_{m-1}x^{m-1} + b_mx^m + 0x^{m+1} + \ldots + 0x^n$.

Определение 50. *Под суммой многочленов $f(x)$ и $g(x)$ будем понимать такой многочлен $h(x) = c_0+c_1x+\ldots+c_nx^n$, коэффициенты которого равны сумме коэффициентов многочленов $f(x)$ и $g(x)$ при одинаковых степенях неизвестного, т.е. $c_i = a_i + b_i$, $i = 0, 1, \ldots, n$. При этом будем записывать $h(x) = f(x) + g(x)$.*

Задача 24. *$(F[x], +)$ — группа.*

Определение 51. *Под произведением $f(x)\cdot g(x)$ будем понимать такой многочлен $h(x) = d_0+d_ix+\ldots+d_{n+m-1}x^{n+m-1}+d_{n+m}x^{n+m}$, коэффициенты которого вычисляются по формуле $d_i = \sum\limits_{k+l=i} a_kb_l$, $i = 0, 1, \ldots, n + m$.*

Задача 25. *$(F[x], +, \cdot)$ — коммутативное кольцо.*

11.1 Основная теорема алгебры

Пусть F — поле и f — произвольный многочлен над F.

Определение 52. *Поле F называется алгебраически замкнутым, если каждый многочлен из кольца $F[x]$ ракладывается на линейные множители.*

То же самое можно выразить другими словами: поле F алгебраически замкнуто, если неприводимыми над F являются лишь многочлены степени 1 (линейные многочлены). Если любой многочлен $f \in F[x]$ обладает в F по крайней мере одним корнем, то поле F алгебраически замкнуто. Действительно, тогда $f(x) = (x - a)h(x)$, $a \in F$, $h \in F[x]$, но по условию для многочлена h в F тоже существует хотя бы один корень,

т.е. $h(x) = (x-b)r(x)$, $b \in F$, $r \in F[x]$. Продолжая этот процесс, мы придём в конце концов к полному разложению f на линейные множители. Так как f — произвольный многочлен, то поле f удовлетворяет определению алгебраической замкнутости. Хотя и справедливо утверждение о том, что для всякого поля F существует расширение $\overline{F} \supset F$, являющееся алгебраически замкнутым полем (теорема Штейница), на первых порах всё же трудно воспринять не только конструкцию алгебраически замкнутого расширения, но и саму идею такого расширения.

Теорема 30. *(Основная теорема алгебры) Поле комплексных чисел* $\mathbb{C}$ *алгебраически замкнуто.*

Сформулируем ещё раз это фундаментальное утверждение, теперь уже в терминах корней.

Произвольный многочлен $f(x)$ степени $n \geq 1$ с комплексными (или вещественными) коэффициентами имеет ровно n комплексных корней, считаемых со своими кратностями.

11.2 Доказательство Основной теоремы алгебры

Его неалгебраичность начинается с двух вспомогательных утверждений.

1) Каждый комплексный многочлен $f(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$, $n \geq 1$, является непрерывной функцией в любой точке плоскости $\mathbb{C}$ (функция $f : \mathbb{C} \to \mathbb{C}$ непрерывна в точке $z_0 \in \mathbb{C}$, если $\lim\limits_{z \to z_0} f(z) = f(z_0)$; другими словами, для любой окрестности $V(f(z_0))$ найдётся окрестность $U(z_0)$ такая, что при любом $z \in U(z_0)$ будет $f(z) \in V(f(z_0))$.

2) Каждая непрерывная функция $f : K \to \mathbb{R}$ на компакте $K \in \mathbb{C}$ достигает своего минимума в K (компакт — замкнутое ограниченное множество). Заметим, что следовало бы говорить о непрерывности полиномиальной функции $f : \mathbb{C} \to \mathbb{C}$, но мы следуем упрощённому языку, принятому в анализе. Компактом у нас будет круг $|z| \leq r$ некоторого достаточно большого радиуса r, определённого ниже. Тривиальный случай многочлена f со свободным членом $a_0 = 0$ исключается из рассмотрения, поскольку тогда f имеет корень $z_0 = 0$.

Чтобы пояснить геометрически идею доказательства, вообразим себе поверхность в $\mathbb{R}^3$, отвечающую уравнению $w = |f(z)|$: значения z изображаются на горизонтальной плоскости $\mathbb{R}^2$, а значения $|f(z)|$ откладываются вверх, в направлении оси w, перпендикулярной к $\mathbb{R}^2$. Из непрерывности $f(z)$ следует непрерывность функции $|f(z)|$ на всей плоскости

$\mathbb{C}$. Нужно убедиться в том, что хотя бы одной точкой наша поверхность "опирается"на горизонтальную плоскость $\mathbb{R}^2$ ($w = 0$). Последующие рассуждения разобьём на несколько шагов.

Лемма 3. *Существует положительное число $r \in \mathbb{R}$ такое, что $|f(z)| > f(0)$ для всех $z \in \mathbb{C}$ с $|z| > r$.*

Доказательство. Действительно, для $z \neq 0$ имеем $|f(z)| = |z|^n|a_n + g(z^{-1})|$, где $g(u) = a_n u + a_{n-1}u^2 + \ldots + a_1 u^n \in \mathbb{C}[u]$. Из непрерывности g в точке 0 следует существование такого вещественного $\delta > 0$, что $|g(u)| \leq |a_n|/2$ при $|u| < \delta$. Таким образом, $|f(z)| \geq |z|^n(|a_n| - |g(z^{-1})|) \geq \frac{1}{2}|a_n||z|^n$ при $|z| > \delta^{-1}$. Следовательно, осталось выбрать любое вещественное число $r > \delta^{-1}$, для которого было бы выполнено неравенство $|a_n|r^n > 2|a_0|$. □

Следствие 8. *(лемма Коши о минимуме). Для каждого многочлена $f \in \mathbb{C}[z]$ существует $z_0 \in \mathbb{C}$ такое, что $|f(z_0)| = \inf\limits_{z\in\mathbb{C}} |f(z)|$.*

Доказательство. В самом деле, ввиду утверждения 2) непрерывная функция $|f(z)|$ принимает в круге $D_r = \{z \in \mathbb{C} || z| \leq r\}$ минимальное значение, т.е. существует $z_0 \in D_r$ такое, что $|f(z_0)| = \inf\limits_{z\in D_r} |f(z)|$. Но так как $|f(z_0)| \leq |f(0)|$, и по лемме 3 имеет место неравенство $|f(0)| \leq \inf\limits_{z\in\mathbb{C}\setminus D_r} |f(z)|$, то $|f(z_0)| = \inf\limits_{z\in\mathbb{C}} |f(z)|$. □

Лемма 4. *Пусть $k \in \mathbb{N}$, и пусть $h \in \mathbb{C}[z]$ — многочлен с $h(0) \neq 0$. Тогда для каждого $a \in \mathbb{C}^*$ найдётся такое $b \in \mathbb{C}$, что $|a + b^k h(b)| < |a|$.*

Доказательство. Так как многочлен h непрерывен, существует $\delta > 0$ такое, что при $|z| < \delta$ имеет место неравенство $|h(z) - h(0)| < h(0)/2$. Это позволяет нам получить оценку для $a + z^k h(z) = a + h(0)z^k + z^k(h(z) - h(0))$:

$$|a + z^k h(z)| \leq |a + h(0)z^k| + \frac{1}{2}|h(0)||z|^k \tag{11.2}$$

из круга $|z| < \delta$.

Выберем теперь комплексное число $b \in \mathbb{C}$, для которого $h(0)b^k = -ta$, $0 < t < 1$ (ниже на вещественное число t будут наложены дополнительные ограничения). В качестве b достаточно взять любой корень степени k из $-tah(0)^{-1} \neq 0$. Получаем $|a + h(0)b^k = (1-t)|a|$ и $\frac{1}{2}|h(0)||b|^k = t|a|/2$, что в соединении с (11.2) приведет к нужному неравенству, коль скоро $|b| < \delta$. Мы обеспечим выполнение этого условия, наложив на $t = -h(0)a^{-1}b^k$ ограничение $t < |h(0)a^{-1}|\delta^k$. Итак, подставив в (11.2) значение $z = b$, $|b| < \delta$, получаем окончательно $|a + b^k h(b)| \leq (1-t)|a| + \frac{1}{2}t|a| = (1 - \frac{1}{2}t)|a| < |a|$. □

Следствие 9. *(лемма Даламбера—Аргана). Пусть $f(z)$ — многочлен положительной степени над $\mathbb{C}$. Тогда каждой точке $c \in \mathbb{C}$ такой, что $f(c) \neq 0$, отвечает точка $c' \in \mathbb{C}$, для которой $|f(c')| < |f(c)|$.*

Доказательство. Для доказательства многочлен $f(z+c)$, подобно $f(z)$ не являющийся константой, разложим по степеням z: $f(z+c) = f(c) + b_k z^k + b_{k+1} z^{k+1} + \ldots + b_n z^n$, $b_k \neq 0$. Другими словами, $f(z+c) = f(c) + z^k h(z)$, где $h(z) = b_k + b_{k+1} z + \ldots + b_n z^{n-k}$, $h(0) \neq 0$. Подставив в формулировку леммы 4 значение $a = f(c) \neq 0$, мы можем утверждать существование такого $b \in \mathbb{C}$, что при $c' = b + c$ будет выполнено требуемое неравенство $|f(c')| = |f(b+c)| = |f(c) + b^k h(b)| < f(c)$. □

Геометрический смысл: если на поверхности $w = f(z)$ взята точка, расположенная строго выше плоскости $w = 0$, то обязательно найдётся другая точка на поверхности с более низким расположением.

Окончание доказательства основной теоремы (теоремы 30). Согласно следствию леммы 3 существует такая точка $z_0 \in \mathbb{C}$, что $|f(z_0)| \leq |f(z)|$ для всех $z \in \mathbb{C}$. Если $|f(z_0)| \neq 0$, то, как утверждает следствие леммы 4, найдётся такая точка $z_0^{'} \in \mathbb{C}$, что $|f(z_0^{'})| < |f(z_0)|$ — противоречие.

11.3 Другое доказательство

Определение 53. *Функция $f : \mathbb{C} \to \mathbb{C}$, $f = u(z) + iv(z)$ — голоморфная, если удовлетворяет условиям (Коши-Римана) $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$, $\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$.*

Теорема 31. *(Интегральная формула Коши) Пусть U — открытая область в $\mathbb{C}$, и $f : U \to \mathbb{C}$ — голоморфная функция, а диск $D = \{z || z - z_0| \leq r\}$ содержится в U. Пусть C — граница D. Тогда $\forall n \in \mathbb{N}$ внутри D верно:*

$$f(a) = (1/(2\pi i)) \oint_C f(z)/(z-a) dz$$

где интеграл берется по контуру, проходимому в положительном направлении (против часовой стрелки).

Доказательство. Можно показать (Теорема Грина), что интеграл по C равен интегралу по произвольно малой окружности, ограничивающей a. Так как $f(z)$ непрерывна, можно выбрать такую окружность, внутри которой $f(z)$ произвольно близка к $f(a)$. С другой стороны

$$\oint_C 1/(z-a)\, dz$$

вдоль любой окружности C с центром в a равен $2\pi i$. (Прямое вычисление

$$z = a + \varepsilon e^{it},$$

где $0 \leq t \leq 2\pi$ и ε — радиус окружности.) Тогда, при $\varepsilon \to 0$ получаем

$$|\frac{1}{2\pi i} \oint\limits_C f(z)/(z-a)dz - f(a)| \leq$$

$$\leq \frac{1}{2\pi i} \oint\limits_C \frac{|f(z)-f(a)|}{z-a} dz \to 0.$$

□

Теорема 32. *Любая голоморфная функция — аналитическая*

Доказательство. Пусть f дифференцируема внутри диска D с центром в $a \in \mathbb{C}$. Пусть $z \in D$. Пусть C — положительно ориентированный контур окружности с центром в a, лежащий внутри D но доальше от a чем z.

$$f(z) = (1/(2\pi i)) \int\limits_C f(w)/(w-z)\, dw =$$

$$= (1/2\pi i) \int\limits_C 1/(w-a) \cdot (w-a)/(w-z) f(w)\, dw =$$

$$= (1/(2\pi i)) \int\limits_C (1/(w-a)) \cdot ((w-a)/((w-a)-(z-a))) f(w)\, dw =$$

$$= (1/(2\pi i)) \int\limits_C (1/(w-a)) \cdot (1/(1-(z-a)/(w-a))) f(w)\, dw =$$

$$= (1/(2\pi i)) \int\limits_C (1/(w-a)) \cdot (\sum_{n=0}^{\infty} ((z-a)/(w-a))^n) f(w)\, dw =$$

$$= \sum_{n=0}^{\infty} (1/(2\pi i) \int\limits_C ((z-a)^n/(w-a)^{n+1}) f(w)\, dw.$$

□

Теорема 33. *(Лиувилль) Функция, аналитическая на всей комплексной плоскости и не имеющая особенностей на бесконечности, есть константа.*

Доказательство. Рассмотрим разложение f в ряд Тэйлора:

$$f(z) = \sum_{k=0}^{\infty} a_k z^k,$$

где

$$a_k = \frac{f^{(k)}}{k!} = 1/(2\pi i) \oint\limits_{C_r} f(\zeta)/\zeta^{k+1} d\zeta,$$

а C_r окружность радиуса r с центром в 0. Тогда

$$|a_k| \leq \frac{1}{2\pi} \oint\limits_{C_r} \frac{|f(\zeta)|}{|\zeta^{k+1}|} d\zeta \leq \frac{1}{2\pi} \oint\limits_{C_r} \frac{M}{r^{k+1}} d\zeta \leq \frac{M}{r^k},$$

где во втором тождестве использовано предположение о том, что $|f(z)| \leq M \ \forall z \in \mathbb{C}$. Эти интегралы не зависят от r. То есть, при $r \to \infty$ $a_k = 0$ $\forall k \geq 1$. Следовательно, $f(z) = a_0$. □

Посему, функция, обратная многочлену должна иметь хоть один полюс на комплексной плоскости, а, соответственно, многочлен имеет хоть один корень.

Глава 12

Расширения полей

12.1 Конечные и алгебраические расширения

Определение 54. *Пусть F — поле. Если F — подполе поля E, то мы говорим также, что E есть расширение поля F. Мы можем рассматривать E как векторное пространство над F, и мы говорим, что E — конечное илн бесконечное расширение F, в зависимости от того, конечна или бесконечна размерность этого векторного пространства.*

Определение 55. *Пусть F — подполе поля E. Элемент α из E называется алгебраическим над F, если в F существуют элементы $a_0, \ldots, a_n$ ($n \geq 1$), не все равные 0 и такие, что*

$$a_0 + a_1\alpha + \ldots + a_n\alpha^n = 0$$

Для алгебраического элемента $\alpha \neq 0$ мы всегда можем найти такие элементы a_i в предыдущем равенстве, что $a_0 \neq 0$ (сокращая на подходящую степень a).

Пусть X — переменная над F. Можно также сказать, что элемент α алгебраичен над F, если гомоморфизм $F[X] \to E$, тождественный на F и переводящий X в α, имеет ненулевое ядро. В таком случае это ядро будет главным идеалом, порожденным одним многочленом $p(X)$, относительно которого мы можем предполагать, что его старший коэффициент равен 1. Имеет место изоморфизм $F[X]/(p(x)) \approx F[\alpha]$, и так как кольцо $F[\alpha]$ целостное, то $p(X)$ неприводим. Если $p(X)$ нормализован условием, что его старший коэффициент равен 1, то $p(X)$ однозначно определяется элементом α и будет называться неприводимым многочленом элемента α над F. Иногда мы будем обозначать его через $\mathrm{Irr}(\alpha, F, X)$.

Определение 56. *Расширение E поля F называется алгебраическим, если всякий элемент из E алгебраичен над F.*

Предложение 11. *Всякое конечное расширение E поля F алгебраично над F.*

Доказательство. Пусть $\alpha \in E$, $\alpha \neq 0$. Степени α $1, \alpha, \alpha^2, \ldots, \alpha^n$ не могут быть линейно независимы над F для всех целых положительных n, иначе размерность E над F была бы бесконечна. Линейное соотношение между этими степенями показывает, что элемент α алгебраичен над F. □

Заметим, что утверждение, обратное предложению 11, не верно: существуют бесконечные алгебраические расширения. Если E — расширение поля F, то мы обозначаем символом $[E : F]$ размерность E как векторного пространства над F. Будем называть $[E : F]$ степенью E над F. Она может быть бесконечной.

Предложение 12. *Пусть k — поле и $F \subset E$ — расширения k. Тогда $[E : k] = [E : F][F : k]$. Если $\{x_i\}_{i \in I}$ — базис поля F над k и $\{y_j\}_{j \in J}$ — базис поля E над F, то $\{x_i y_j\}_{(i,j) \in I \times J}$ будет базисом поля E над k.*

Доказательство. Пусть $z \in E$. По предположению существуют элементы $\alpha_j \in F$, почти все равные нулю и такие, что $z = \sum\limits_{j \in J} \alpha_j y_j$. Для каждого $j \in J$ существуют элементы $b_{ij} \in k$, из которых почти все равны 0, такие, что $\alpha_j = \sum\limits_{i \in I} b_{ji} x_i$, следовательно, $z = \sum j \in J \sum\limits_{i \in I} b_{ji} x_i y_j$. Это означает, что $\{x_i y_j\}$ является семейством образующих для E над k. Мы должны показать, что оно линейно независимо. Пусть $\{c_{ij}\}$ — семейство элементов из k, почти все из которых равны 0, такое, что $\sum j \in J \sum\limits_{i \in I} c_{ij} x_i y_j = 0$. Тогда для каждого j $\sum i \in I c_{ij} x_i = 0$, поскольку элементы y_j линейно независимы над F. Наконец, $c_{ij} = 0$ для всякого i, так как $\{x_i\}$ — базис поля F над k, что и доказывает наше предложение. □

Следствие 10. *Расширение $E \supset F \supset k$ поля k конечно в том и только в том случае, если E конечно над F и F конечно над k.*

Назовем башней полей последовательность расширений $F_1 \subset F_2 \subset \ldots \subset F_n$. Для конечности башни необходимо и достаточно, чтобы каждый ее этаж был конечен.

Пусть k — поле, E — его расширение и $\alpha \in E$. Мы обозначаем через $k(\alpha)$ наименьшее подполе в E, содержащее k и α. Оно состоит из всех дробей $f(\alpha)/g(\alpha)$, где f, g — многочлены с коэффициентами в k и $g(\alpha) \neq 0$.

Предложение 13. *Пусть элемент α алгебраичен над k. Тогда $k(\alpha) = k[\alpha]$ и поле $k(\alpha)$ конечно над k. Степень $[k(\alpha) : k]$ равна степени многочлена* $\mathrm{Irr}(a, k, X)$.

Доказательство. Пусть $p(X) = \mathrm{Irr}(\alpha, k, X)$. Пусть многочлен $f(X) \in k[X]$ таков, что $f(\alpha) \neq 0$. Тогда $f(X)$ не делится на $p(X)$ и, следовательно, существуют многочлены $g(X)$, $h(X) \in k[X]$, такие, что $g(X)p(X) + h(X)f(X) = 1$. Отсюда мы получаем, что $h(\alpha)f(\alpha) = l$ и, значит, $f(\alpha)$ обратим в $k[\alpha]$. Следовательно, $k[\alpha]$ не только кольцо, но и поле, а потому должно быть равно $k(\alpha)$. Пусть $d = \deg p(X)$. Степени $1, \alpha, \ldots, \alpha^{d-1}$ линейно независимы над k; действительно, предположим, что $a_0 + a_1\alpha + \ldots + a_{d-1}\alpha^{d-1} = 0$, где $a_i \in k$, причем не все $a_i = 0$. Положим $g(X) = a_0 + a_1X + \ldots + a_{d-1}X^{d-1}$. Тогда $g \neq 0$ и $g(\alpha) = 0$.

Следовательно, $g(X)$ делится на $p(X)$ — противоречие. Наконец, пусть $f(\alpha) \in k[\alpha]$, где $f(X) \in k[X]$. Существуют многочлены $q(X)$, $r(X) \in k[X]$, такие, что $\deg r < d$ и $f(X) = q(X)p(X) + r(X)$. Тогда $f(\alpha) = r(\alpha)$ и мы видим, что $1, \alpha, \ldots, \alpha^{d-1}$ порождают $k[\alpha]$ как векторное пространство над k. Это доказывает наше предложение. □

Пусть E, F — расширения поля k. Если E и F содержатся в не- котором поле L, то мы обозначаем через EF наименьшее подполе в L, содержащее и E, и F, и называем его композитом E и F в L,

Литература

[1] Кострикин А. И., Введение в алгебру.Ч.1-3, Москва, ФИЗМАТЛИТ, 2004.

[2] Артамонов В.А., Лекции по алгебре, МГУ, 2004.

[3] S. S. Adams, Introduction to Algebraic Coding Theory, 2008